AI TR
Tsinghua

人工智能之认知图谱
Research Report of Cognitive Graph

2020 年第 6 期



清华大学人工智能研究院

北京智源人工智能研究院

清华—中国工程院知识智能联合研究中心

阿里集团—新零售智能引擎事业群

2020 年 08 月

VISIT…

LANZAROTE
Caliente.com

扫码免费领取资料



房中术(www.nblax520.com)专注于男性增大增粗增长、阳痿、早泄、壮阳、延时、强肾、回春、健身。女性缩阴、丰胸、减肥、化妆、瑜伽、保养、产后修复、盆底肌锻炼。两性健康,夫妻按摩,房中术,性姿势,性技巧,性知识等

更多免费教程：英语学习，技能提升，PS 教学，投资赚钱，音乐教程，口才教学，情商提升，风水教学，心理学，摄影知识，幼儿学习，书法学习，记忆力提升等等......

扫码加微信lax988988领取资料



==

全站课程下载

课程不断增加

本站现资源容量已超 10T

入群联系 QQ：167520299 或添加微信：1131084518（备注 学习）

阳痿早泄训练
皇室洗髓功视频教学
女人驻颜术
泡妞约炮万元课程
足疗养 SPA 教材
玉蛋功
马氏回春功
房中术张丰川
哲龙全套视频
增大盼你增大
国际男优训练
亚当德永早泄训练
洗髓功真人内部
皇室养生绝学道家洗髓功
【铁牛人会员课】男人必备技能，理论讲解
实战高清视频
随意控制射精锻炼 视频+图片+文字
价值 1440 元第一性学名著<素女经房中养生宝典视频>12 部
洗髓功修炼方法视频教学
陈见玉蛋功视频教学 女性缩阴锻炼
男性自然增大增长指南
强性健肾保健操 1-4
道家强肾系统锻炼功法
马氏回春功
12 堂课，全面掌握男性健康问题 让你重燃自信
联系微信：1131084518

1、东方性经
2、印度 17 式
3、口交技巧 3 部
4、港台性姿势 3 部
5、365 性姿势 6 部
6、泰国性爱密经 17 式
7、花花公子性技巧 6 部
8、阁楼艳星性技巧 7 部
9、古今鸳鸯秘谱全集 7 部
10、夫妻爱侣情趣瑜伽 2 部
11、古代宫廷性保健系列 14 部
12、汉唐宋元明清春宫图真人
13. 柔软性爱宝典 日本 9800 课
14. 李熙墨 3999 全套课
15. 妖精性爱课 2888
16. 李银河全套性课
17. 领统统性课
18. 德勇男性篇
19. 德勇男性篇
20. 缓慢性爱
21. 亚当多体位搭配篇
22. 亚当多体位结合篇
23. 德勇克服早泄讲座练习
24. 德勇以女性为中心得爱抚
25. 加藤鹰接吻爱抚舌技
26. 加藤鹰指技
27. 加藤鹰四十八手入门
28. 佐藤潮吹教学
29. 佐藤男人体能锻炼+保健品介绍
30. 佐藤男人早泄对抗训练
31. 阿拉伯延时训练

32. 田渊正浩秘籍
33. 异性性快感集中训练教学
34. 自我愉悦锻炼密宗
35. 铁牛全套延时训练课
36. Pc 机锻炼真人视频教学
37. 印度性经全集 8 部
38. 21 世纪性爱指南
39. 香蕉大叔男女训练馆全套
40. 中美真人性治疗教学+理论
41. 女性闺房秘术
42. 幸福玛利亚性课
43. 陈见如何释放性魅力征服
44. 性爱技巧讲座全套
45. 性爱秘籍全套
46. 性爱误区讲座
47. 性病讲解大全
48. 性博士讲座合集
49. 性健康和性高潮合集
50. 性教育讲座合集
51. 性能力课堂合集
52. 性生活问题解析合集
53. 意外怀孕和避孕处理课堂
54. 性感地带探索
55. 性技巧讲座
56. 性健康与性卫生讲座
57. 性生活专家答疑
58. 性心理与性道德合集
59. 性爱宝典合集
60. 性爱技巧合集
61. 完美性爱演示
62. 完美性爱技术讲解
更多精品等你来解锁哦...........

人工智能之认知图谱

Research Report of Cognitive Graph

AMiner

2020 年第 6 期

清华大学人工智能研究院

北京智源人工智能研究院

清华—中国工程院知识智能联合研究中心

阿里集团—新零售智能引擎事业群

2020 年 08 月

如果群里报告过期

请加微信联系我索取最新

1、每日微信群内分享7+最新重磅报告;

2、每日分享当日华尔街日报、金融时报

3、如果看到群里报告过期了，请扫码联系

4、行研报告均为公开版，权利归原作者所有，仅分发做内部学习

扫一扫二维码后台回复 加群

加入“研究报告”微信报告群。



摘要

认知图谱（Cognitive Graph）旨在结合认知心理学、脑科学和人类知识等，研发融合知识图谱、认知推理、逻辑表达的新一代认知引擎，支持大规模知识的表示、获取、推理与计算的基础理论和方法，实现人工智能从感知智能向认知智能的演进，建立可解释、鲁棒性的第三代人工智能。

本报告围绕认知图谱的概念内涵、关键技术、人才研究、应用场景、发展趋势等方面展开深入研究，主要内容包括：

一、认知图谱基本概念、产生历程、机遇与挑战。详细介绍了认知图谱的产生背景和基本概念，以及在演化过程中出现的代表性事件，并总结了认知图谱在当前环境下的机遇与挑战。

二、认知图谱基础理论和技术研究现状分析。根据认知图谱的概念，分别对知识图谱、认知推理、逻辑表达等领域的概念背景、发展历程、关键技术、应用、问题与挑战、未来研究方向等方面进行了详细介绍和深入分析。

三、认知图谱领域人才现状分析。基于 AMiner 平台提供的论文和学者大数据，从学者分布、学术水平、国际合作、学者流动等维度，对国内外相关研究学者和机构进行了对比分析，总结中国科研学者队伍建设过程中的弱势环节和问题，并提出对策建议。

四、认知图谱典型应用场景分析。首先以阿里巴巴电商平台为例，详细介绍了认知图谱如何赋能电商平台的搜索、推荐等核心业务。然后介绍了认知图谱在智慧城市、司法行业、金融行业、安防行业、精准分析、智慧搜索、智能推荐、智能解释、自然人机交互等行业技术的应用场景和案例。

最后分析了认知图谱相关技术研究发展趋势和创新热点，以及中国的专利数据和国家自然科学基金支持情况，并展望了认知图谱未来发展方向。

AMiner

目录

1 概述篇..2
1.1 认知图谱概念................................2
1.2 认知图谱产生历程............................4
1.3 认知图谱机遇与挑战..........................8
2 技术篇..12
2.1 知识图谱....................................12
2.1.1 知识图谱概念..............................12
2.1.2 知识图谱发展历程..........................14
2.1.3 知识图谱关键技术..........................15
2.1.4 知识图谱应用..............................54
2.1.5 知识图谱研究问题与挑战....................55
2.1.6 知识图谱未来研究方向......................56
2.2 认知推理....................................58
2.2.1 知识图谱推理概念..........................58
2.2.2 知识图谱推理关键技术......................59
2.2.3 知识图谱推理应用..........................74
2.2.4 知识图谱推理研究问题与挑战................76
2.2.5 知识图谱推理未来研究方向..................77
2.3 逻辑表达....................................80
2.3.1 自然语言生成概念..........................80
2.3.2 自然语言生成关键技术......................82
2.3.3 自然语言生成应用..........................91
2.3.4 自然语言生成研究问题与挑战................92
2.3.5 自然语言生成未来研究方向..................94
2.4 论文主题分析................................95
2.5 经典论文解读................................99
2.6 技术情报挖掘................................104
3 人才篇..110

AMiner

3.1 学者情况概览......110
3.1.1学者分布地图......110
3.1.2学术水平分析......112
3.1.3国际合作分析......115
3.1.4学者流动情况......117
3.2 代表性学者画像......119
3.2.1国外代表性学者......121
3.2.2国内代表性学者......131
3.3 中国学者问题与对策......140
4 应用篇......144
4.1 电商平台......144
4.1.1认知推荐......145
4.1.2基础数据层......145
4.1.3推理引擎层......147
4.1.4用户交互的文本和视觉智能......154
4.2 其他应用场景......167
4.2.1行业应用......167
4.2.2技术应用......169
5 趋势篇......176
5.1 技术研究发展趋势......176
5.2 技术研究创新热点......177
5.3 中国专利数据情况......179
5.4 国家自然科学基金支持情况......180
6 总结与展望......184
参考文献......187
附录 1 认知图谱相关的关键词列表......201
附录 2 代表性期刊和会议列表......202
附录 3 国家自然科学基金 NSFC 项目......210
AMiner

图表目录

图 1 人工智能发展的几个阶段 2
图 2 双通道理论框架... 3
图 3 认知图谱的演化历程..................................... 5
图 4 知识图谱样例... 13
图 5 知识图谱发展历程....................................... 14
图 6 TransE 模型的简单示例 17
图 7 TransH 模型的简单示例 18
图 8 TransR 模型的简单示例 19
图 9 KG2E 模型的示例 21
图 10 传统模型和 TransG 模型比较............................ 22
图 11 RESCAL 模型的简单图解 23
图 12 DistMult 模型的简单图解 24
图 13 HOlE 模型的简单图解 25
图 14 SME 模型的神经网络结构 26
图 15 NTN 模型的神经网络结构 27
图 16 MLP 模型的神经网络结构 27
图 17 NAM 模型的神经网络结构 28
图 18 知识图谱的体系架构.................................... 30
图 19 实体抽取样例.. 31
图 20 实体链接案例.. 38
图 21 RDF 图示例：电影知识图谱 45
图 22 属性图示例：电影知识图谱.............................. 46
图 23 属性表存储方案示例.................................... 49
图 24 Neo4j 中顶点和边记录的物理存储结构 52
图 25 VS*树.. 54
图 26 描述逻辑的语义表...................................... 61

AMiner

图 27 描述逻辑与 OWL 词汇的对应表................................61
图 28 Tableaux 的运算规则..62
图 29 Tableaux 的相关工具简介....................................62
图 30 逻辑编程改写的相关工具简介................................63
图 31 基于一阶查询重写方法的处理流程............................64
图 32 产生式规则系统的执行流程..................................65
图 33 产生式规则方法的相关工具..................................65
图 34 使用逻辑规则用于知识图谱推理任务的示例....................66
图 35 RESCAL 模型...70
图 36 自然语言产生框架..81
图 37 询问天气场景中的句子模板..................................82
图 38 询问天气场景中的词汇模板..................................83
图 39 ELMo 模型架构...84
图 40 Transformer 模型架构......................................86
图 41 GPT 模型结构..87
图 42 BERT 模型的架构...89
图 43 XLNet 的 Attention Mask 原理机制...........................91
图 44 LDA 结构图..96
图 45 认知图谱相关技术概览.....................................105
图 46 全球学者分布地图...111
图 47 中国学者分布地图...112
图 48 中国与其他国家的论文合作情况.............................117
图 49 全球学者的流动情况.......................................118
图 50 中国学者的流动情况.......................................119
图 51 AMiner 平台学者画像示例..................................120
图 52 阿里巴巴认知智能计算平台.................................145
图 53 阿里巴巴数据生态...146
图 54 ATRank 的网络结构..148
AMiner

图 55 从用户行为学习解离化表征模型实现架构总览图..............153
图 56 背包颜色维度示例图..154
图 57 背包大小维度示例图..154
图 58 KOBE 模型的基本框架..156
图 59 商品个性化推荐示例图......................................157
图 60 多模态表征学习框架..158
图 61 买家秀视频推荐的两个实际应用场景..........................161
图 62 商品关联属性信息图..162
图 63 Gavotte 模型结构图..163
图 64 GLA 模型结构图..165
图 65 部分实验结果示例图..167
图 66 M-Recnet 和 M-LiveBot 模型的结果对比图......................167
图 67 认知图谱领域的技术研究发展趋势............................177
图 68 认知图谱领域的技术创新热点词云图..........................178
图 69 中国历年的专利数量分布（2010-2019 年）......................179
图 70 2010-2019 年中国专利数量 TOP 10 机构........................180
图 71 认知图谱相关领域国家自然科学基金项目支持历年分布情况....181
图 72 认知图谱相关领域国家自然科学基金项目支持数量 TOP 10 机构统计
....182

表 1 认知智能和感知智能的异同点....................................5
表 2 知识图谱产品统计...13
表 3 三元组表示例...49
表 4 认知图谱领域论文主题分布.....................................96
表 5 h-index TOP 5000 全球学者的国家统计..........................111
表 6 h-index TOP 5000 中国学者的省市统计..........................112
表 7 论文总被引频次排名前 10 的国家...............................113
表 8 论文总被引频次排名前 10 的全球机构...........................114

AMiner

表 9 论文总被引频次排名前 10 的中国机构........................115
表 10 合作论文数量排名前 10 的国家列表........................115
表 11 学术指标说明..119
表 12 认知图谱研究热点子领域的代表性学者的学术指标统计........178
表 13 认知图谱相关领域国家自然科学基金项目分类情况（2010-2020 年）
....180
表 14 认知图谱相关领域的关键词列表............................201
表 15 认知图谱领域代表性期刊和会议列表........................202
表 16 认知图谱领域国家自然科学基金支持的相关项目（2010-2020 年）
....210

AMiner

1　概述篇



1 概述篇

1.1 认知图谱概念

大规模常识知识库与基于认知的逻辑推理是人工智能发展的瓶颈问题。1968 年图灵奖获得者 Edward Feigenbaum 研发出世界首个专家系统 DENDRAL，并随后在第五届国际人工智能会议上提出，将知识融入计算机系统是解决只有领域专家才能解决的复杂问题的关键；1999 年互联网发明人、图灵奖获得者 Tim Berners-Lee 爵士提出语义网的概念，核心理念是用知识表示互联网，建立常识知识；2019 年图灵奖获得者 Yoshua Bengio 在 NeurIPS 大会上的主题报告中指出深度学习应该从感知为主向基于认知的逻辑推理和知识表达方向发展，这个思想和清华大学张钹院士近期提出的第三代人工智能思路不谋而合。总的来说人工智能的发展经历了表示、计算到感知两个阶段，下一个阶段的核心是认知（见图 1）。近期图灵奖获得者 Manuel Blum 夫妇提出意识 AI（意识智能）的思想，这是一个既经典又全新的概念和思路。核心的理念就是构造一个新型的可用数学建模、可计算的机器认知/意识模型。

AMiner



图 1 人工智能发展的几个阶段

阿里巴巴达摩院发布“2020 十大科技趋势”报告中提到，人工智能已经在“听、说、看”等感知智能领域达到或超越了人类水准，但在需要外部知识、逻辑推理或者领域迁移的认知智能领域还处于初级阶段[1]。复旦大学肖仰华教授在报告“知识图谱与认知智能”中描述了，让机器具备认知智能具体体现在机器能够理解

数据、理解语言进而理解现实世界的能力，体现在机器能够解释数据、解释过程进而解释现象的能力，体现在推理、规划等等一系列人类所独有的思考认知能力上，需要去解决推理、规划、联想、创作等复杂任务[2]。如何实现认知智能，阿里巴巴新零售智能引擎事业群总监杨红霞在 CNCC2018 会议上的报告“*Extremely large scale cognitive graph representation in practice*”详细介绍了阿里电商认知图谱，以更好地认知用户需求为目标，将助力搜索推荐等从基于行为的方式迈向基于行为与语义融合的认知智能时代[3]。

清华大学唐杰教授在“人工智能下一个十年”报告[4]中，结合认知科学和计算机理论，给出了一个实现认知智能的可行思路：认知图谱=知识图谱+认知推理+逻辑表达，希望利用知识表示、推理和决策，包括人的认知来解决复杂问题。这个思路的基本思想是结合认知科学中的双通道理论，在人脑的认知系统中存在两个系统：System 1 和 System 2，如图 2 所示。System 1 是一个直觉系统，它可以通过人对相关信息的一个直觉匹配寻找答案，它是非常快速、简单的；而 System 2 是一个分析系统，它通过一定的推理、逻辑找到答案。比如，针对以下问题：“找到一个 2003 年在洛杉矶的 Quality 咖啡馆拍过电影的导演”。System 1 首先找到相关的影片，然后用 System 2 来做决策，如果是标准答案，就结束整个推理的过程。如果不是标准答案，而相应的信息又有用，就把它作为一个有用信息提供给 System 1，System 1 继续做知识的扩展，System 2 再来做决策，直到最终找到答案。图灵奖获得者 Bengio 在 NIPS 2019 大会的 Keynote 也提到，System 1 到 System 2 的认知是深度学习未来发展的重要方向。

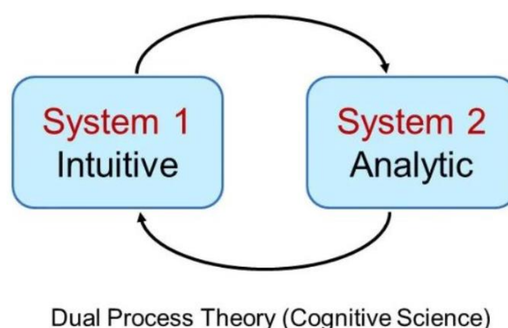


图 2 双通道理论框架

AMiner

到目前为止，认知图谱在学术界和工业界还没有一个统一定义。百度百科给出的认知图谱定义如下：认知图谱（Cognitive Graph）旨在结合认知心理学、脑科学和人类知识等，研发知识图谱、认知推理、逻辑表达的新一代认知引擎，实现人工智能从感知智能向认知智能的演进。认知图谱是计算机科学的一个研究分支，它企图了解智能的实质，并实现感知智能系统到认知智能系统的重大技术突破[5]。

清华大学唐杰教授的“认知图谱—人工智能的下一个瑰宝”文章中，认知图谱的核心是以实现融合知识驱动和数据驱动相结合的知识表示和推理的认知引擎为目标，研究支持鲁棒可解释人工智能的大规模知识的表示、获取、推理与计算的基础理论和方法；建设包含语言知识、常识知识、世界知识、认知知识的大规模知识图谱以及典型行业知识库，建成知识计算服务平台。

清华大学丁铭博士的“从知识图谱到认知图谱：历史、发展与展望”文章中，认知图谱被解释为“基于原始文本数据，针对特定问题情境，使用强大的机器学习模型动态构建的，节点带有上下文语义信息的知识图谱”。

让机器具备认知智能，其核心就是让机器具备理解和解释能力。这种能力的实现与大规模、结构化的背景知识是密不可分的。认知图谱是实现机器认知智能的使能器，一方面通过建立从数据到认知图谱中实体、概念、关系的映射，让机器理解数据的本质；另一方面利用认知图谱中实体、概念和关系来解释现实世界中事物和现象，让机器解释现象的本质。

1.2 认知图谱产生历程

从现阶段人工智能的发展来看，随着计算力的不断发展，储存手段的不断升级，具有快速计算和记忆存储能力的计算智能可以说已经实现；而随着移动互联网普及，大数据、云计算等技术发展，更多非结构化数据的价值被重视和挖掘，语音、图像、视频、触点等与感知相关的感知智能也在快速发展；在计算智能和感知智能发展基础上，人工智能正在向能够分析、思考、理解、判断等认知智能延伸，真正的智能化解决方案已经显现端倪。下表展示了认知智能和感知智能的

AMiner

异同。

表 1 认知智能和感知智能的异同点

差异点	认知智能	感知智能
特点	强调认知、理解	以人控为主，由人告知机器如何行动
要求能力	有学习、推理能力，能通过分析做出恰当决策，为人们提供参考	接受人的训练、培养
与外界关系	与人、环境之间有互动，增加人类智慧	没有互相的反馈，主要由人控制，根据人类的需求而工作
衡量标准	没有统一标准，遇到具体问题具体分析	有像图灵测试这样的衡量标准
计算行为	以大规模数据为背景，模拟人的思考行为	统计分析大规模数据

认知图谱结合大规模、结构化的背景知识，利用知识表示、推理和决策，通过了解智能的实质，让机器具备理解和解释事物本质的能力，是实现感知智能系统到认知智能系统的重大技术突破一种有效手段。图 3 展示了认知图谱演化过程中出现的代表性事件。

AMiner

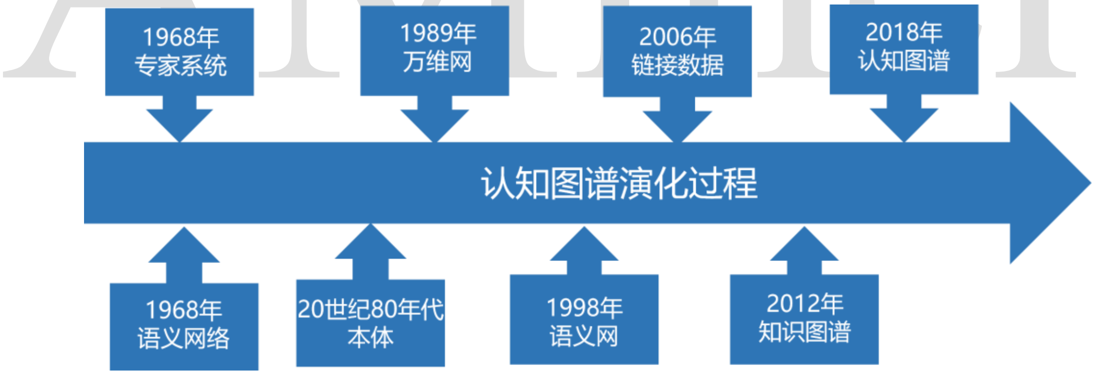


图 3 认知图谱的演化历程

认知图谱的历程发展可以追溯到语义网络（Semantic Network）。1968 年，Quillian 在文献 *Semantic Memory in M. Minsky*[6]研究人类长期记忆模型时，描述了人类长期记忆的一般结构模型，认为记忆由概念之间的联系来实现并存储在复杂的网络中，并基于此提出了语义网络的概念。语义网络是通过概念及其语义关系来表达知识的一种网络图。随着互联网的发展，语义网络有了新的应用场景——语义网（Semantic Web）。

1968 年，美国斯坦福大学的 Edward Feigenbaum（图灵奖获得者）等人在文献 *Heuristic DENDRAL: A program for generating explanatory hypotheses*[7]中，提出了首个专家系统 DENDRAL，旨在帮助化学家判断某待定物质的分子结构。专家系统是一个具有大量的专门知识与经验的程序系统，它应用人工智能技术和计算机技术，根据某领域一个或多个专家提供的知识和经验，进行推理和判断，模拟人类专家的决策过程，以便解决那些需要人类专家处理的复杂问题。

自二十世纪 70 年代中期以来，人工智能（简称 AI）领域的研究人员认识到，知识的获取乃是构建强大 AI 系统的关键所在，他们借助于来自哲学本体论的灵感，认为通过将本体（Ontology）创建成为计算模型，可以成就特定类型的自动推理。二十世纪 80 年代，AI 领域研究人员开始将本体应用到信息科学领域。1991 年美国斯坦福大学的 Thomas R. Gruber 在文献 *The role of common ontology in achieving sharable, reusable knowledge bases*[8]中介绍了一种构建可共享、可重用知识库的策略，其中本体作为知识耦合构造角色，在其中发挥核心作用。本体是（特定领域）信息组织的一种形式，是领域知识规范的抽象和描述，是表达、共享、重用知识的方法。本体是知识体系构建的关键技术，通过对知识建模，使计算机能够识别人类知识，相当于知识图谱的模式架构（Schema）。1984 年，美国微电子与计算机技术公司的 Douglas Lenat 设立的 Cyc 项目是最早的本体知识库，旨在。1994 年，Cyc 项目从该公司独立出去，并以此为基础成立了 Cycorp 公司。

1989 年，英国的计算机科学家 Tim Berners-Lee（万维网联盟 W3C 主管）发明了万维网（World Wide Web）。万维网通过超文本标记语言（HTML）把信息组织成为图文并茂的超文本，利用链接从一个站点跳到另个站点，以此摆脱以前查询工具只能按特定路径一步步地查找信息的限制。2017 年，Tim Berners-Lee 因“发明万维网、第一个浏览器和使万维网得以扩展的基本协议和算法”获得 2016 年度的图灵奖。

1998 年，Tim Berners-Lee 再次在文献 *Semantic Web Road Map*[9]中提出语义网（Semantic Web）。语义网的核心是：通过给万维网上的文档（如：HTML 文档、XML 文档）添加能够被计算机所理解的语义“元数据”（Meta Data），从而使

AMiner

整个互联网成为一个通用的信息交换媒介。语义网是能够根据语义进行判断的智能网络，它不但能够理解词语和概念，而且还能够理解它们之间的逻辑关系，可以使交流变得更有效率和价值。

2006 年，Tim Berners-Lee 提出链接数据（Linked Data）的概念，数据不仅仅发布于语义网中，而要建立起数据之间的链接从而形成一张巨大的链接数据网。链接数据提出的目的是构建一张计算机能理解的语义数据网络，而不仅仅是人能读懂的文档网络，以便于在此之上构建更智能的应用。链接数据也可以是开放数据，在这种情况下通常称为链接开放数据（Linked open data, LOD）。

2012 年，Google 的 Singhal 等人在博客 *Introducing the Knowledge Graph: things, not strings*[10]中介绍了知识图谱（Knowledge Graph）的概念。知识图谱以语义网络的结构化方式描述客观世界中概念、实体、事件以及它们之间的关系，相对于传统的本体和语义网络而言，实体覆盖率更高，语义关系也更加复杂而全面。Google 使用语义检索从多种数据来源收集信息构建知识图谱，以提高 Google 搜索的质量。

AMiner

2018 年，阿里巴巴新零售智能引擎事业群总监杨红霞在 CNCC2018 会议上的报告"*Extremely large scale cognitive graph representation in practice*"详细介绍了阿里电商认知图谱，以更好地认知用户需求为目标，将助力搜索推荐等从基于行为的方式迈向基于行为与语义融合的认知智能时代[3]。2020 年，清华大学唐杰教授在"人工智能下一个十年"报告[4]中，结合认知科学中的双通道理论和计算机理论，给出了一个实现认知智能的可行思路：认知图谱=知识图谱+认知推理+逻辑表达，希望利用知识表示、推理和决策，包括人的认知来解决复杂问题。

人工智能作为新一轮产业变革的核心驱动力和引领未来发展的战略技术，引起了世界各国的高度重视，并取得了快速发展。然而大规模常识知识库与基于认知的逻辑推理是人工智能发展的瓶颈问题。认知图谱以实现融合知识驱动和数据驱动相结合的知识表示和推理的认知引擎为目标，是实现鲁棒可解释人工智能的关键技术手段。

1.3 认知图谱机遇与挑战

尽管人工智能依靠深度学习和机器学习技术的进步取得了巨大的进展，例如，AlphaGo 通过自我强化学习击败了人类顶尖的围棋选手，但人工智能在很多方面，如语言理解、视觉场景理解、决策分析等，仍然举步维艰。当前的人工智能识别做的只是比对，缺少信息进入大脑之后的加工、理解、思考等步骤，因此仅仅停留在“感知”，而并非“认知”。要让人工智能有类似大脑的活动，走到认知阶段，需要让它掌握知识、进行推理。机器必须要掌握大量的知识，特别是常识知识才能实现真正类人的智能。

目前的智能系统在感知方面已经达到甚至超越人类水平，但在鲁棒性、可解释性、安全可靠等方面还存在很多不足。感知智能技术存在的缺陷，包括但不限于：

（1）模型鲁棒性差，难以与准确性共生。为了揭示深度神经网络模型的鲁棒性和准确性之间的关系，来自 IBM 研究院，加州大学戴维斯分校，麻省理工学院以及京东 AI 研究院的研究人员，系统性地度量了 18 个被学术界和工业界广泛接受并使用的 ImageNet 深度神经网络模型，如 AlexNet、VGG Nets、Inception Nets、ResNets、DenseNets、MobileNets、NASNets 等的鲁棒性。该研究发现：1）准确度越高的模型的鲁棒性普遍越差，且分类错误率的对数和模型鲁棒性存在线性关系；2）相比于模型的大小，模型的结构对于鲁棒性的影响更大。

（2）模型可解释性差，对于可靠性要求高的任务很难胜任。比如大部分阅读理解方法都只能看做黑盒，从输入问题和文档到输出答案文本块，问题与答案中间的练习部分处于黑盒中心。用户无法看到从问题到答案的推理过程，从而无法验证答案的对错。要使认知图谱具有可解释性，需要向用户展示推理路径或者子图，每个推理节点上的支撑事实，或是用于对比的其他可能答案和推理路径。

（3）缺乏积累知识的能力，也没能和人类已有的知识体系进行很好的关联，缺乏可靠的推理方法。比如对于多跳问答的问题，问题和答案间拥有多少相似词汇已经不能作为答案检索的重要依据。检索时需要有知识支撑，通过知识理解问

AMiner

题与答案之间的联系。若知识的涵盖度不够，或检索时仅仅理解局部片段，而非整个文档，缺乏在知识层面上的推理能力，就会出现短视检索的问题。

人类社会已经进入智能时代，智能时代的社会发展催生了大量的智能化应用，智能化应用对机器的认知智能化水平提出了前所未有的要求。繁杂的应用场景、深度的知识应用、密集的专家知识给认知图谱的研究带来了巨大挑战[11]。近年虽然人工智能取得快速发展，但如何将深度学习与大规模常识知识结合起来，实现认知推理与逻辑表达还面临很大挑战。

（1） 繁杂的应用场景。由于企业与领域应用呈现出鲜明的复杂特性，因此构建企业知识图谱面临的场景是“小而杂”，比如员工报销的审批流程涉及员工基本信息、天气、交通、财务报销制度等。而计算机喜欢用通用模型处理问题，如果需要针对不同场景定制模型，给数据处理、模型学习等方面都会带来巨大挑战。

（2） 深度的知识应用。通过构建数据驱动的统计模型在用户行为数据丰富、应用模式简单的搜索与推荐等场景中应用效果良好，但是难以解决面向复杂决策的智能运维、医疗诊断、司法研判等应用场景的实际问题。这些应用场景需要知识驱动的智能系统，通过大规模背景知识支撑模型学习、推理和决策。

（3） 密集的专家知识。由于大部分领域应用需要密集的专家知识，比如故障排查、医疗问诊等。而专家知识的获取渠道通常是行业从业人员，难以直接从开放数据源中抽取，导致专家知识的数量稀少，无法满足现阶段领域应用的需求。

（4） 复杂的数据资源。如今时代虽然产生了海量数据，但是大数据环境下数据的分布、异构、动态、碎片化和低质等特征给知识工程和知识服务提出了新挑战，也使得高度依赖数据规模和质量的机器学习和深度学习模型在部分场景问题中难以达到预期的解决效果。针对既需要从感知角度学习数据的分布表示，又需要从认知角度解释数据语义的需求，构建新一代开放常识知识图谱和研发认知推理核心技术成为实现下一代人工智能技术突破的关键。

（5） 自适应的知识推理。人工智能时代多样化的数据特征和任务需求，对研究大规模多粒度自学习的知识推理技术提出了新挑战：基于深度学习与逻辑推

AMiner

理相结合的大规模多粒度知识推理，基于本体、规则与深度学习相结合的大规模知识推理，以及实现亿级三元组和万级规则的快速推理；基于时空特性的知识演化模型和推理规则的自学习，研制知识、推理、逻辑的演化系统，能够根据外界反馈实时对知识进行更新，实现推理规则的自学习和逻辑表达的自学习。

总之，无论是在具备智能理解能力还是安全性方面，以感知智能技术为主的人工智能与人类智能相去甚远。随着数据红利消耗殆尽，以深度学习为代表的感知智能遇到天花板。认知智能将是未来一段时期内人工智能发展的焦点，是进一步释放人工智能产能的关键，而结合知识图谱、认知推理、逻辑表达的认知图谱是实现认知智能技术突破的关键，不仅让机器理解数据本质，还可以让机器解释现象本质。

AMiner

2　技术篇



2 技术篇

基于前述内容，认知图谱的研究内容主要包括知识图谱、认知推理和逻辑表达，因此本章节重点介绍以上 3 个研究方向的概念、发展历程、关键技术、问题与挑战、未来研究方向等信息，旨在为读者了解认知图谱领域的基础、应用研究的代表性成果、研究动向和进展提供信息窗口。

2.1 知识图谱

知识图谱被认为是从感知智能通往认知智能的重要基石。一个很简单的原因就是，没有知识的机器不可能实现认知智能。图灵奖获得者，知识工程创始人 Edward Feigenbaum 曾经提到："Knowledge is the power in AI system"。张钹院士也提到，"没有知识的 AI 不是真正的 AI"。从感知到认知的跨越过程中，构建大规模高质量知识图谱是一个重要环节，当人工智能可以更结构化地表示理解人类知识并进行互联，才有可能让机器真正实现推理、联想等认知功能。

2.1.1 知识图谱概念

AMiner

知识图谱（Knowledge Graph, KG）将互联网的信息表达成更接近人类认知世界的形式，提供了一种更好地组织、管理和理解互联网海量信息的能力。知识图谱给互联网语义搜索带来了活力，同时也在智能问答、大数据分析与决策中显示出强大威力，已经成为互联网基于知识的智能服务基础。知识图谱与大数据和深度学习一起，成为推动人工智能发展的核心驱动力之一。

知识图谱以结构化的方式描述客观世界中概念、实体、事件以及它们之间的关系，本质上是一种语义网络，网络节点表示实体或者概念，边表示实体或者概念之间的关联关系。其中，概念是指人们在认识世界过程中形成的对客观事物的概念化表示，如国家、城市、人物等；实体是客观世界中的具体事物，如中国、北京等。属性也是一种实体，用于描述事物的内在信息，比如中国的面积、人口等。关系描述概念、实体之间客观存在的关联，如首都描述了城市和国家的关系，国家和城市之间概念和子概念的关系等。图 4 展示了知识图谱样例。

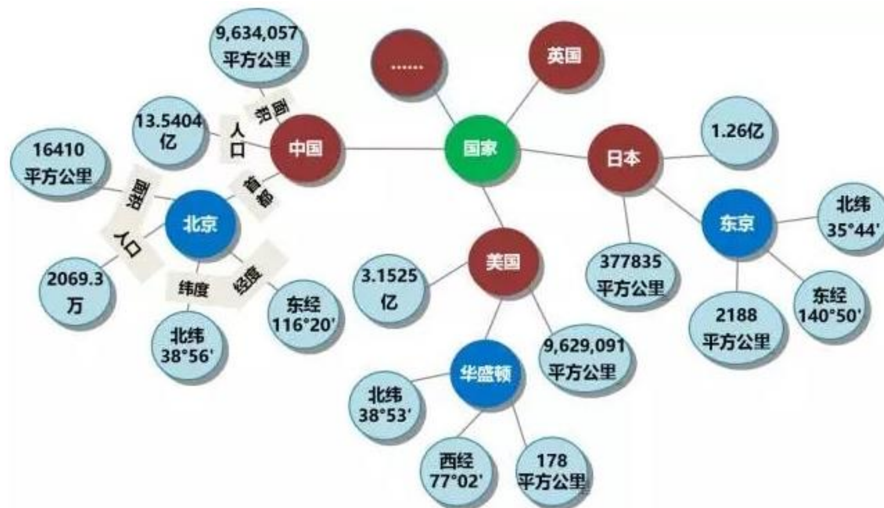


图 4 知识图谱样例

虽然知识图谱的概念由谷歌在 2012 年提出，但早在此之前，出现了大量的相关产品。尤其在 2006 年，Berners-Lee 提出了数据链接（linked data）的思想，通过对 URI（uniform resource identifier）、RDF（resource description framework）和 OWL（web ontology language）等技术标准的推广和完善，掀起了一场语义网络研究热潮，基于语义网络的研究成果，衍生了很多知识库成果。表 2 介绍了一些著名的通用领域知识图谱产品的规模。我国对中文知识图谱的研究虽然起步较晚，但取得了很多有价值的研究成果。比如早期由中国科学院计算机语言信息中心发布的知网（HowNet），清华大学构建的 XLORE 是第一个大规模中英文跨语言知识图谱，复旦大学推出的中文知识图谱项目，上海交通大学发布的中文知识图谱研究平台 zhishi.me 等等。中文知识图谱对中文信息的处理和检索具有重要的应用价值，近年来吸引了大量的研究。

AMiner

表 2 知识图谱产品统计

知识图谱名称	实体数量	事实三元组数量	访问网址
WordNe	117,597	207,016	https://wordnet.princeton.edu
OpenCyc	47,000	306,000	https://www.cyc.com/opencyc/
Cyc	~250,000	~2,200,000	https://www.cyc.com
YAGO	1,056,638	~5,000,000	http://www.mpii.mpg.de/~suchanek/yago
DBpedia	~1,950,000	~103,000,000	https://wiki.dbpedia.org/develop/datasets

Freebase	-	~125,000,000	https://developers.google.com/freebase/
NELL	-	242,453	http://rtw.ml.cmu.edu/rtw/
Wikidata	14,449,300	-	https://www.wikidata.org/wiki
Probase IsA	12,501,527	85,101,174	https://concept.research.microsoft.com/Home/Download
Google KG	> 500 million	> 3.5 billion	https://developers.google.com/knowledge-graph
XLORE	2,466,956	16,284,901	https://xlore.org/

2.1.2 知识图谱发展历程

知识图谱实质上是一种图形化知识表示形式。知识表示在逻辑和人工智能领域经历了漫长的发展历史。图形化知识表示的思想最早可以追溯到 1956 年 Richens[12]提出的语义网概念，而符号逻辑知识可以追溯到 1959 年文献[13]提出的一般问题解决方法。知识库首先被用于基于知识的推理和问题解决系统。MYCIN[14]是最著名的基于规则的医学诊断专家系统之一，其中知识库约有 600 条规则。后来，人类知识表示社区推动了基于框架的语言、基于规则的表示和混合表示的发展。大约在这个时期的末期，Cyc 项目[15]开始了，旨在汇集人类的知识。资源描述框架（Resource Description Framework，RDF）[16]和 Web 本体语言（OWL）[17]相继发布，成为语义 Web[18]的重要标准。然后，许多开放知识库或本体被发布，如 WordNet、DBpedia、YAGO 和 Freebase。Stokman 和 Vries[19]在 1988 年提出了一个使用图进行知识结构化表示的现代概念。然而，2012 年谷歌提出利用知识图谱，提升搜索引擎的质量，知识图谱的概念才得到了广泛关注，当时谷歌利用 Knowledge vault[20]（一种知识融合框架）来构建大规模的知识图谱。图 5 展示了知识图谱的发展路线[21]。

AMiner

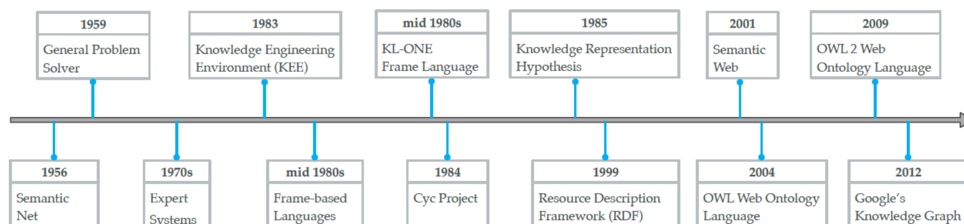


图 5 知识图谱发展历程

从上图可以看出，知识图谱的发展历程大致可以归纳为：从 20 世纪五六十年代提出语义网络，20 世纪 70 年代出现了语义网络和一阶逻辑谓词的关系，20 世纪 80 年代开始完善语义网络理论（语义推理），20 世纪 90 年代产生了概念关系建模，21 世纪出现了新的应用场景（语义 web）：通过 W3C 的标准拓展 web，实现数据在不同应用中的共享和重用，用户上传各种图结构的数据（W3C 的标准 RDF）形成连接数据，最后 2012 年谷歌提出知识图谱（Knowledge Graph）概念。谷歌于 2012 年 5 月推出谷歌知识图谱，旨在增强其搜索引擎的搜索结果，标志着大规模知识在互联网语义搜索中的成功应用。

通过以上介绍，可以发现知识图谱的前身（比如 WordNet、Cyc 项目、HowNet）主要通过专家制定规则或者提供专业知识进行构建的。这类知识资源的质量虽然有保证，但是规模较小、领域局限性较强。后来为了扩大知识规模，出现了基于群体智能的构建方法，具体是利用维基百科、百度百科、互动百科等百科类数据源中的半结构化信息（比如信息框），比如 DBpedia、Freebase 以构建本体的形式对维基百科中的知识条目进行组织，YAGO 融合了 WordNet 中良好的概念层次结构和维基百科中的大量知识条目数据。清华大学和上海交通大学通过利用维基百科、百度百科、互动百科，建立了大规模知识图谱 XLORE 和 Zhishi.me[22]。其中，XLORE 是融合中英文维基、法语维基和百度百科，对百科知识进行结构化和跨语言链接构建的多语言知识图谱，是中英文知识规模较平衡的大规模多语言知识图谱。大数据时代，互联网上存在海量的信息资源，为了提高知识图谱的覆盖度，满足智能化信息发展的需求，从互联网数据自动获取知识是建立可持续发展知识图谱的发展趋势。面向互联网上大规模、碎片化、异构的数据环境，学术界和工业界研发了一系列利用机器学习和信息抽取技术自动抽取知识的方法。比如华盛顿大学图灵中心的 KnowItAll[23]、TextRunner[24]，卡内基梅隆大学的 NELL[25]等。

2.1.3 知识图谱关键技术

知识图谱技术是指在建立知识图谱中使用的技术，是融合认知计算、知识表示与推理、信息检索与抽取、自然语言处理与语义 Web、数据挖掘与机器学习等技术的交叉研究。知识图谱研究，一方面探索从互联网语言资源中获取知识的理

AMiner

论和方法；另一方面促进知识驱动的语言理解研究。随着大数据时代的到来，研究从大数据中挖掘隐含的知识理论与方法，将大数据转化为知识，增强对互联网资源的内容理解，将促进当代信息处理技术从信息服务向知识服务转变。具体地，知识图谱技术包括知识图谱表示、知识图谱构建和知识图谱存储等三方面的研究内容[26]。

（1） 知识图谱表示

知识图谱表示（Knowledge Graph Representation），又称知识图谱嵌入（Knowledge Graph Embedding），关键思想是将知识图谱中的实体和关系映射到连续的向量空间中，以便简化操作，同时保留知识图谱的固有结构。实体和关系嵌入表示有利于多种任务的执行，包括知识图谱补全、关系抽取、实体分类和实体解析等。

目前大部分研究方法仅根据观察到的事实（实体和关系的三元组形式）来执行嵌入任务。知识图谱嵌入技术包含三个阶段：（1）实体和关系表示；（2）定义一个评分函数；（3）学习实体和关系表示。对于给定的知识图谱，首先将实体和关系表示为连续的空间向量，并对每个事实定义一个评分函数来度量其合理性。通过最大化观察到的事实的总可信度来获得实体和关系的嵌入表示。在整个过程中，所学习的实体和关系嵌入表示只需要在每个单独的事实中兼容，因此对于下游任务可能没有足够的预测性。因此，越来越多的研究人员开始进一步利用其他类型的信息，如实体类型、关系路径、文本描述，甚至是逻辑规则，来学习更多的预测性嵌入表示。

现有的知识图谱嵌入技术主要分为两类：平移距离模型（Translational Distance Models）、语义匹配模型（Semantic Matching Models）[27]。平移距离模型利用基于距离的评分函数，对两个实体之间的关系执行平移操作后，通过两个实体之间的距离来度量事实的合理性。语义匹配模型利用基于相似性的评分函数，通过匹配实体的潜在语义和向量空间表示中包含的关系来度量事实的可信性。

下面对平移距离模型的代表性研究方法（TransE 模型及其扩展形式、高斯嵌入模型（Gaussian Embeddings））、语义匹配模型的代表性研究方法（RESCAL 模

AMiner

型及其扩展形式、基于神经网络的匹配模型）展开介绍。

- **平移距离模型—TransE 模型及其扩展形式**

本节选择了几个具有代表性的平移距离模型，包括 TransE、TransH、TransD、TransR，分别展开介绍。

✧ **TransE 模型**

Mikolov 等人于 2013 年提出了 word2vec 词表示学习模型和工具包[28]。利用该模型，Mikolov 等人发现词向量空间存在平移不变现象。受到该平移不变现象的启发，Border 等人提出了 TransE 模型[29]，将知识图谱中的关系看作实体间的某种平移向量。对于每个事实三元组（h, r, t），TransE 模型将实体和关系表示在同一向量空间中，把关系向量 r 看作头实体向量 h 和尾实体向量 t 之间的平移，即 h+r≈t。比如：对于给定的 2 个事实（姜文, 导演, 邪不压正）和（冯小刚, 导演, 芳华），除了可以得到：姜文+导演≈邪不压正、冯小刚+导演≈芳华，还可以通过平移不变性得到：邪不压正-姜文≈芳华-冯小刚，即得到两个事实相同的关系的向量表示。我们也可以将 r 看作从 h 到 t 的翻译，因此 TransE 也被称为翻译模型，如图 6 所示，对于每一个三元组（h, r ,t）TransE 希望：h+r≈t，评分函数被定义为（h+r）和 t 之间的负距离，如公式（1）所示。

$$f_r（h，t）=-\|h+r-t\|_{1/2}$$

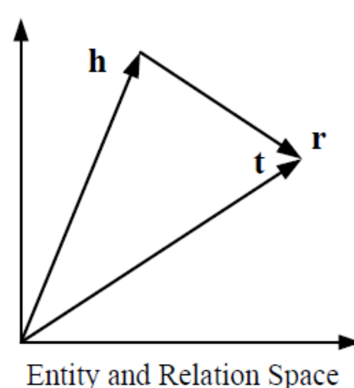


图 6 TransE 模型的简单示例

TransE 模型的参数较少，计算的复杂度明显降低，并且在大规模稀疏知识图谱上也同样具有较好的性能与可扩展性。但是 TransE 模型不能用在处理复杂关

AMiner

系上（比如一对多、多对一、多对多等 3 种关系类型）。原因如下：以一对多为例，对于给定的事实，以姜文拍的民国三部曲电影为例，即《让子弹飞》、《一步之遥》和《邪不压正》。可以得到三个事实三元组即（姜文，导演，让子弹飞）、（姜文，导演，一步之遥）和（姜文，导演，邪不压正）。按照上面对于 TransE 模型的介绍，可以得到，让子弹飞≈一步之遥≈邪不压正，但实际上这三部电影是不同的实体，应该用不同的向量来表示。

✧ **TransH 模型**

为了解决 TransE 模型在处理一对多、多对一、多对多复杂关系时的局限性，TransH 模型[30]提出让一个实体在不同的关系下拥有不同的向量表示。在这种策略下，虽然《让子弹飞》、《一步之遥》和《邪不压正》的向量表示在“导演”关系的情况下非常相似，但是在其他关系下，它们的向量形式相差较大。

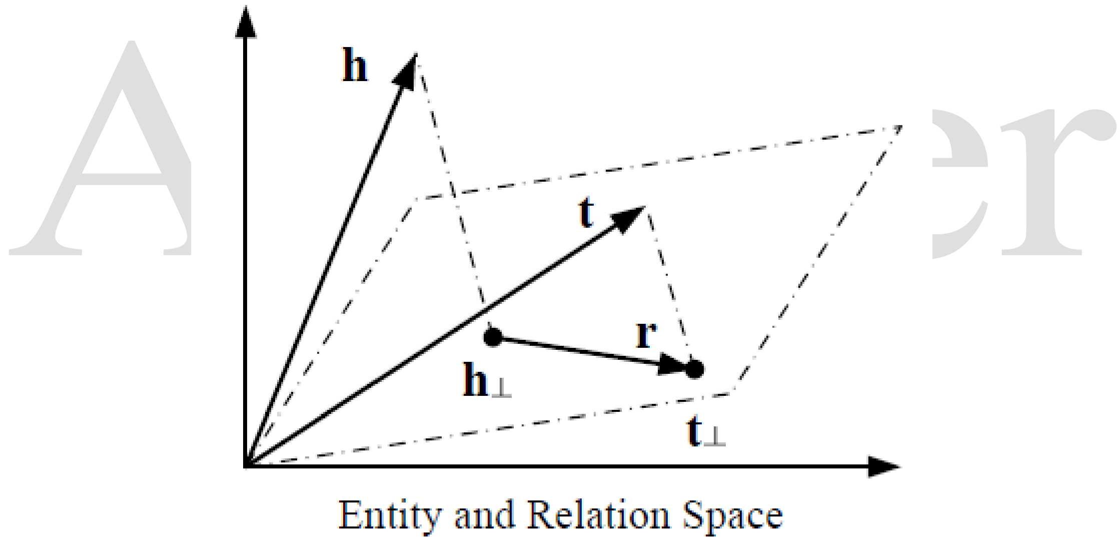


图 7 TransH 模型的简单示例

基于以上观点，TransH 模型引入了关系超平面，如图 7 所示。TransH 模型把实体建模为向量形式，但是对于关系 r，TransH 模型同时使用平移向量 r 和超平面的法向量 wr 来表示它。对于一个三元组（h, r, t），TransH 首先将头实体向量 h 和尾实体向量 t，沿着法向量 wr 投影到关系 r 对应的超平面，产生如下公式：

$$h_{\perp}=h - w_r^{\top} h w_r,\quad t_{\perp}=t - w_r^{\top} t w_r.$$

如果（h, r, t）成立，则假定投影由超平面上的关系向量 r 连接，使得误差较小，即 h⊥+r ≈ t⊥。因此，得分函数（如下所示）和 TransE 模型中的一样。通过引入映射到关系超平面的机制，TransH 实现了实体在不同关系中的不同角色。需要注意的是，由于关系 r 可能存在无限个超平面，TransH 简单地令 r 与 wr 近似正交来选取某一个超平面。TransH 使不同的实体在不同的关系下拥有了不同的表示形式，但由于实体向量被投影到了关系的语义空间中，故它们具有相同的维度。

$$f_r（h, t）=-\|h_\perp+r-t_\perp\|_2^2$$

✧ **TransR 模型**

虽然 TransH 模型使每个实体在不同关系下拥有了不同的表示，它仍然假设实体和关系处于相同的语义空间中，这一定程度上限制了 TransH 的表示能力。TransR 模型[31]则认为，一个实体是多种属性的综合体，不同关系关注实体的不同属性。TransR 认为不同的关系拥有不同的语义空间。对每个三元组（h, r, t），首先应将头实体和尾实体的向量表示投影到对应的关系空间中（如下公式所示）；

$$h_\perp=M_r h,\ \ t_\perp=M_r t$$

$M_r \in R^{k\times d}$是实体空间到关系空间的投影矩阵。然后再建立从头实体到尾实体的平移关系，评分函数如下所示。

$$f_r（h, t）= -\|h_\perp+r\ -t_\perp\|_2^2$$

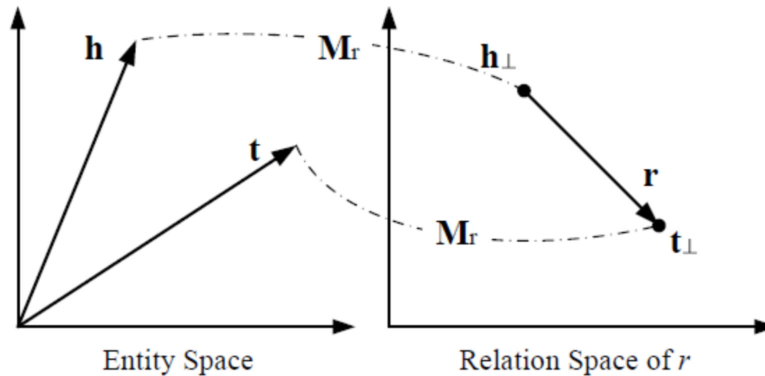


图 8 TransR 模型的简单示例

如图 8 展示了 TransR 模型的简单示例。虽然 TransR 在复杂关系建模方面功

AMiner

能强大，但是它存在以下缺点：

1) 在同一个关系下，头、尾实体共享相同的投影矩阵。然而，一个关系的头、尾实体的类型或属性可能差异巨大。例如，对于三元组（美国，总统，奥巴马），美国和奥巴马的类型完全不同，一个是国家，一个是人物。

2) 从实体空间到关系空间的投影是实体和关系之间的交互过程，因此 TransR 让投影矩阵仅与关系有关是不合理的。

3) 由于 TransR 模型为每个关系引入了一个投影矩阵，每个关系需要 O（dk）个参数。从而使得 TransR 模型比 TransE 和 TransH 模型复杂，因为 TransE 和 TransH 模型只是将关系建模为向量表示形式，并且每个关系只需要 O（d）个参数。

✧ **TransD 模型**

TransD 模型通过将投影矩阵分解为两个向量，简化了 TransR 模型。具体来说，针对事实三元组（h, r, t），根据实体和关系向量表示形式 h, t∈ R^d和 r∈ R^k ，TransD 模型引入了额外的映射向量 w_h, w_t∈ R^d 和 w_r ∈ R^k,，两个投影矩阵分别定义如下：

$$M_r^1=w_r w_h^\top+I，M_r^2=w_r w_t^\top+I$$

这两个投影矩阵分别应用到头实体和尾实体，来获得它们的投影，如下所示。

$$h_\perp = M_r^1 h，t_\perp = M_r^2 t$$

针对投影的实体，评分函数与 TransR 一样。因为 TransD 需要 O（nd+mk）个参数，而 TransR 需要 O（nd+mdk）个参数，因此 TransD 模型效率更高。

● **平移距离模型—高斯嵌入模型**

本节对 KG2E 和 TransG 等两个具有代表性的高斯嵌入模型的实现原理展开介绍。

✧ **KG2E 模型**

文献[32]认为知识图谱中的实体和关系的语义本身存在不确定性，而以往的知识图谱嵌入研究成果都忽略了这个因素，因此该文献提出了 KG2E 模型，使用高斯分布来表示实体和关系。其中高斯分布的均值表示的是实体或关系在语义空

AMiner

间中的中心位置，而高斯分布的协方差则表示该实体或关系的不确定度。图 9 展示了 KG2E 模型示例，每个圆圈代表不同实体与关系的表示，它们分别与“比尔·克林顿”构成三元组，其中圆圈大小表示的是不同实体或关系的不确定度，可以看到“国籍”的不确定度远远大于其他关系。

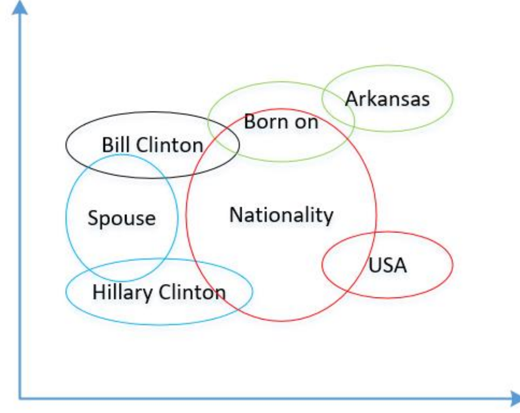


图 9 KG2E 模型的示例

AMiner

KG2E 模型将实体和关系表示为从多变量高斯分布中抽取的随机向量：

$$h \sim N(\mu_h, \Sigma_h),$$

$$t \sim N(\mu_t, \Sigma_t),$$

$$r \sim N(\mu_r, \Sigma_r)$$

其中，μ_h、μ_t、μ_r是均值向量，Σ_h、Σ_t、Σ_r代表协方差矩阵。然后，受到平移假设的启发，KG2E 模型通过测量 t-h 和 r 这两个随机向量之间的距离来为一个事实评分，即 N（μ_t-μ_h，Σ_t+Σ_h）和 N（μ_r，Σ_r）这两个分布。通过 2 种方法来进行测量。

一种是通过 KL 散度（KL 距离）来进行测量，公式如下：

$$f_r(h, t) = -\int N_x(\mu_t-\mu_h, \Sigma_t+\Sigma_h) \ln \frac{N_x(\mu_t-\mu_h, \Sigma_t+\Sigma_h)}{N_x(\mu_t, \Sigma_r)} dx$$

$$\propto -\mathrm{tr}(\Sigma_r^{-1}(\Sigma_h+\Sigma_t)) - \mu^{\top}\Sigma_r^{-1}\mu - \ln \frac{\det(\Sigma_t)}{\det(\Sigma_h+\Sigma_t)},$$

另一种方法是计算概率的内积即：

$$f_r\ (h,t) = \int N_x\ (\mu_t\text{-}\mu_h,\ \Sigma_t+\Sigma_h).\ N_x\ (\mu_r,\Sigma_r)\ dx$$

$$\propto -\mu^{\top}\Sigma^{-1}\ \mu\text{-}\ \ln\ (\det(\Sigma)).$$

其中 $\mu=\mu_h+\ \mu_r\text{-}\mu_r$，$\sum=\Sigma_h+\Sigma_r+\Sigma_t$利用高斯嵌入。KG2E 可以有效地对知识图谱中实体和关系的不确定性进行建模。

✧ **TransG 模型**

TransG[33]也是对高斯分布的实体进行了建模，即：TransG 提出使用高斯混合模型描述头、尾实体之间的关系。该模型认为，一个关系会对应多种语义，每种语义用一个高斯分布来刻画，即:

$$\mathrm{r}=\sum_i \pi_r^i\,\mu_r^i,\ \ \mu_r^i\sim N\ (\mu_t\text{-}\mu_h,\ (\sigma_h^2+\sigma_t^2)\,\mathbf{I}).$$

这里，μ_r^i是第 i 个语义的嵌入，π_r^i是该语义的权重，I 是单位矩阵。从评分函数公式可以看出，平移距离的混合是由关系的不同语义引入的。这些语义组合可以通过 CRP 从数据中自动学习。

TransG 模型与传统模型的对比如图 10 所示。其中三角形表示正确的尾实体，圆形表示错误的尾实体。图 10（a）中为传统模型示例，由于将关系 r 的所有语义混为一谈，导致错误的实体无法被区分开。而如图 10（b）所示，TransG 模型通过考虑关系 r 的不同语义，形成多个高斯分布，就能够区分出正确和错误实体。

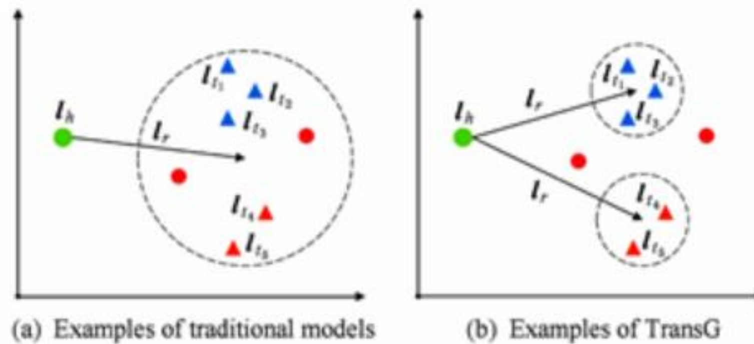


图 10 传统模型和 TransG 模型比较

● **语义匹配模型——RESCAL 模型及其扩展形式**

本节选择三个具有代表性的语义匹配模型分别展开介绍，包括 RESCAL 模型、DistMult 模型和 HolE 模型。

✧ RESCAL 模型

RESCAL（又称双线性模型）[34]通过使用一个向量表示每个实体来获得它的潜在语义。每个关系表示为一个矩阵，该矩阵对潜在因素之间的成对交互作用进行了建模。它把事实（h,r,t）评分函数定义为一个双线性函数如下所示。

$$f_r（h, t）= h^{\top} M_r t = \sum_{i=0}^{d-1} \sum_{j=0}^{d-1} [M_r]_{ij} \cdot [h]_i \cdot [t]_j$$

其中 $h, t \in R^d$表示头实体和尾实体的向量形式，$M_r \in R^{d\times d}$表示关系矩阵。这个分数可以捕获头实体向量 h 和尾实体向量 t 的所有组成部分之间成对的相互作用（见图 11），每个关系需要$R^{d\times d}$个参数。

文献[35]进一步假设对所有的M_r在一组普通的秩为 1 的矩阵上进行分解，即：$M_r = \sum_i \pi_r^i u_i v_i^{\top}$。TATEC 模型[36]不仅对三向交互$h^{\top} M_r t$ 进行建模，还对双向交互（如实体与关系之间的交互）进行建模。评分函数是 f_r（h, t）= $h^{\top} M_r t + h^{\top} r + t^{\top} r + h^{\top} D t$，其中 D 是所有不同关系共享的对角矩阵。

AMiner

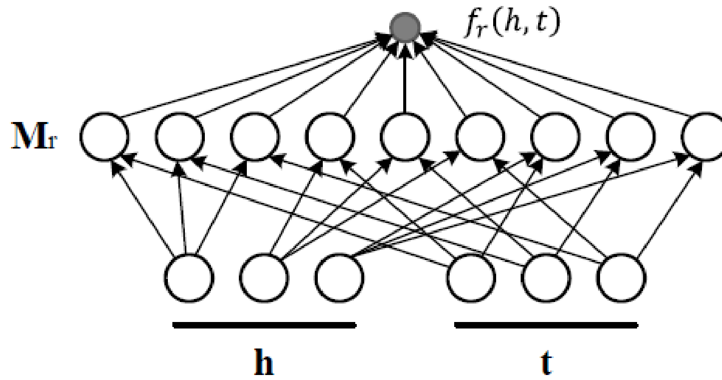


图 11 RESCAL 模型的简单图解

✧ DistMult 模型

DistMult[37]通过将 Mr 限制为对角矩阵来简化 RESCAL。对于每一个关系 r，DistMult 模型都引入一个向量嵌入 r，并要求 Mr=diag（r），评分函数如下所示：

$$f_r（h, t）= h^{\top} \mathrm{diag}(r) t = \sum_{i=0}^{d-1} [r]_i \cdot [h]_i \cdot [t]_i$$

评分函数只捕获同一维度的 h 和 t 分量之间的成对交互作用（如图 12 所示），并将每一个关系的参数数量减少至 O（d）。然而，因为对于任意的 h 和 t，$h^{\top}\mathrm{diag}(r)t = t^{\top}\mathrm{diag}(r)h$都是成立的，这种过度简化的模型只能处理对称的关系，这显然对于通用知识图谱是不能完全适用的。

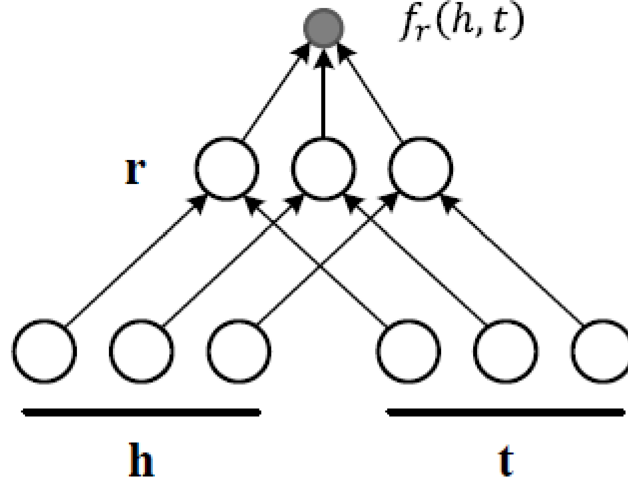


图 12 DistMult模型的简单图解

✧ **HolE 模型**

AMiner

HolE 模型[38]将 RESCAL 模型的表达能力与 DistMult 模型的效率和简单性相结合。它把实体和关系都表示为 Rd 中的向量。给定一个事实（h, r, t），首先使用循环关联操作将实体表示形式组合为$h * t \in R^d$，即：

$$[h * t]_i = \sum_{k=0}^{d-1} [h]_k . [t]_{(k+i) mod\ d}$$

然后将实体组合向量与关系表示形式匹配，以对事实进行评分，即：

$$f_r\ (h, t) = r^{\top}\ (h*t) = \sum_{i=0}^{d-1}[r]_i \sum_{k=0}^{d-1}[h]_k . [t]_{(k+i) mod\ d}$$

循环关联对成对的相互作用进行压缩（如图 13 所示）。因此，HolE 对每个关系只需要 O（d）个参数，因此比 RESCAL 模型更高效。与此同时，因为循环关联是不符合交换律的，即 h*t≠t*h。所以 HolE 能够像 RESCAL 那样建模不对称关系。

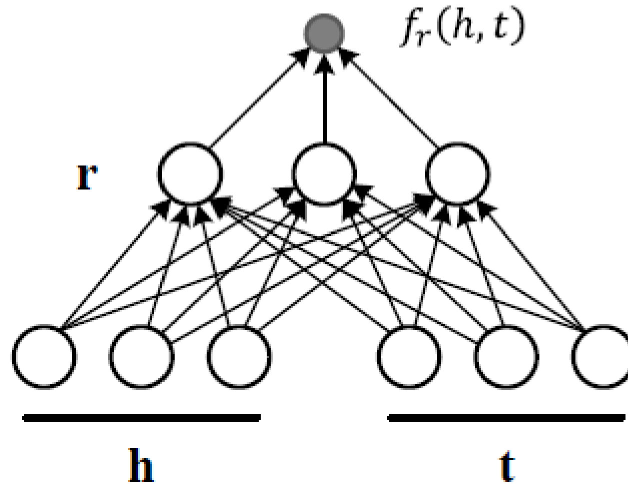


图 13 HOlE模型的简单图解

- **语义匹配模型——基于神经网络的匹配模型**

本节选择 4 个具有代表性的基于神经网络的语义匹配模型（语义匹配能量模型、神经张量网络模型、多层感知机、神经关联模型）的实现原理进行介绍。

✧ **语义匹配能量模型（SME）**

AMiner

语义匹配能量模型（Semantic Matching Energy）[39]采用神经网络结构进行语义匹配。给定一个事实三元组（h, r, t），它首先将实体和关系投影到输入层中的嵌入向量。然后将关系 r 与头实体 h 组合得到g_u（h, r），关系 r 与尾实体 t 组合得到隐藏层中的g_v（t, r）。则该事实三元组的分数最终由它们的点积定义为匹配的g_u 和g_v 。即：

$$f_r(h,t)=g_u(h,r)^{\top}g_v(t,r)$$

SME 有两个版本：线性版本和双线性版本。SME（线性版本）定义为：

$$g_u(h,r)=M_u^1h+M_u^2r+b_u,$$
$$g_v(t,r)=M_v^1h+M_v^2r+b_v,$$

SME（双线性版本）定义为：

$$g_u(h,r)=M_u^1h\circ M_u^2r+b_u,$$
$$g_v(t,r)=M_v^1h\circ M_v^2r+b_v,$$

其中，M_u^1，M_u^2，M_v^1，$M_v^2 \in R^{d\times d}$是权重矩阵，而b_u，$b_v \in R^d$是跨不同关系共享的偏差向量。图 14 展示了 SME 模型的简单神经网络结构。

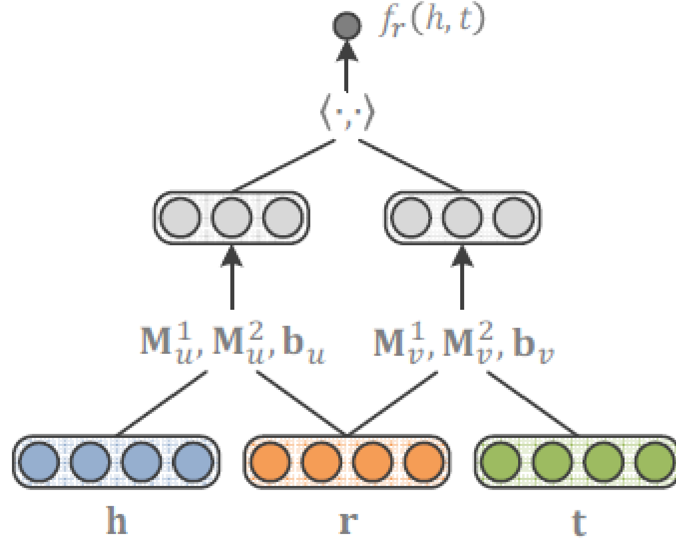


图 14 SME模型的神经网络结构

✧ **神经张量网络模型（NTN）**

AMiner

神经张量网络模型（Neural Tensor Network, NTN）是另外一种神经网络结构。给定一个事实三元组（h, r, t），它首先将实体投影到输入层中的嵌入向量。然后将这两个实体 h，t∈ R^d由关系特有的张量$M_r \in R^{d\times d\times k}$（以及其他参数）组合，并映射到一个非线性隐藏层。最后，一个针对关系的线性输出层给出了评分，即：

$$f_r(h,t)=r^\top \tanh(h^\top M_r t + M_r^1 h + M_r^2 t + b_r)$$

其中M_r^1,$M_r^2 \in R^{k\times d}$和$b_r \in R^k$分别是关系的权重矩阵和偏差向量。双线性张量积$h^\top M_r t$ 得到一个向量，其中第 i 项条目计算为 $h^\top M_r^{[:,:,i]} t$。图 15 给出了 NTN 模型的一个简单说明。通过以下设置M_r=0 和$b_r = 0$，NTN 退化为单层模型（SLM）。尽管 NTN 是迄今为止最具表达能力的模型，但是它针对每个关系需要 $O(d^2k)$个参数，因此不能简单有效地处理大规模的知识图谱。

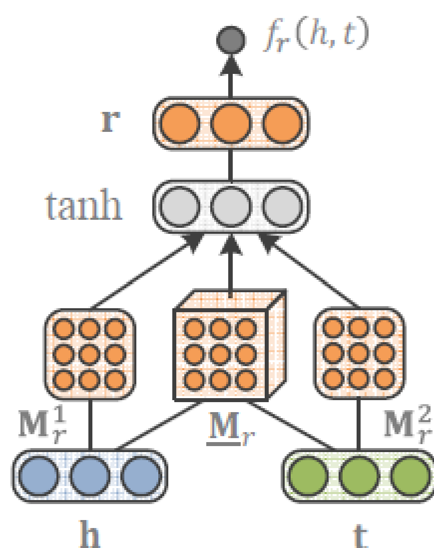


图 15 NTN模型的神经网络结构

✧ 多层感知机（MLP）

多层感知机（Multi-Layer Perceptron, MLP）是一种更简单的方法，在这种方法中，每个关系（以及实体）都是由一个向量组合而成的，如图 16 所示。给定一个事实（h, r, t），将嵌入向量 h、r 和 t 连接在输入层中，并映射到非线性的隐藏层。然后由线性输出层生成分数，即：

$$f_r\ (h,\ t) = w^\top\ tanh\ (M^1 h + M^2 r + M^3 t)$$

其中$M^1, M^2, M^3 \in R^{d\times d}$是第一层的权重，w$\in R^d$是第二层的权重，这些都是在不同的关系中共享的。

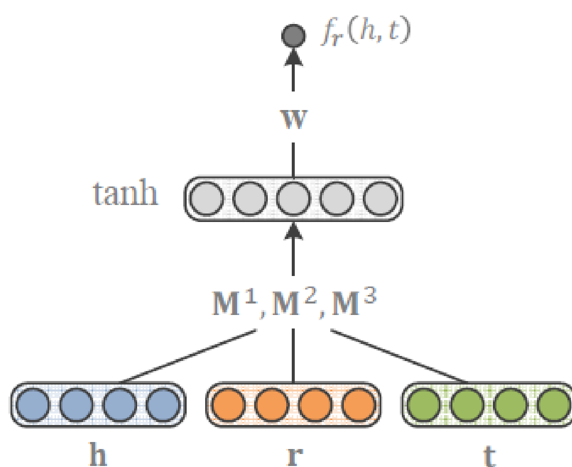


图 16 MLP模型的神经网络结构

AMiner

✧ 神经关联模型（NAM）

神经关联模型（Neural Association Model, NAM）使用“深度”架构进行语义匹配。给定一个事实三元组（h, r, t），它首先将头实体的嵌入向量和输入层中的关系连接起来，从而给出$z^{(0)} = [h; r] \in R^{2d}$。然后$z^{(0)}$输入到一个由 L 个线性隐层组成的深神经网络中，即：

$$a^{(l)} = M^{(l)} z^{(l-1)} + b^{(l)}, \quad l = 1, . . . , L,$$

$$z^{(l)} = ReLU(a^{(l)}), \quad l = 1, . . . , L,$$

其中，$M^{(l)}$ 和$b^{(l)}$ 分别表示第 1 层的权重矩阵和偏差。在前馈过程之后，通过匹配最后一个隐藏层的输出和尾实体的嵌入向量来给出分数，即：

$$f_r(h, t) = \mathbf{t}^{\top} \mathbf{z}^{(L)}$$

图 17 展示了 NAM 模型的简单说明，它有一个更复杂的版本，将关系嵌入 r 连接到网络中的所有隐藏层。

AMiner

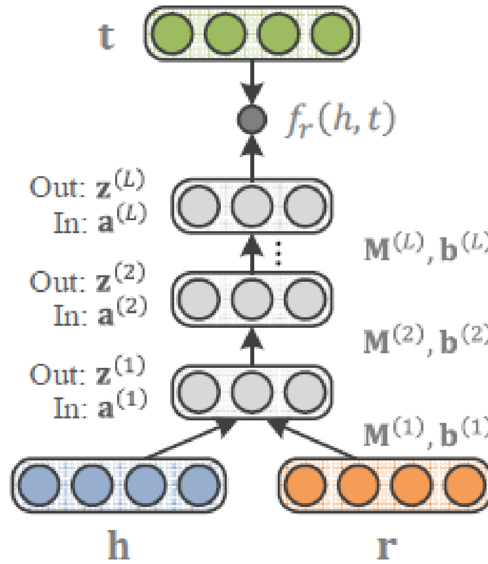


图 17 NAM模型的神经网络结构

（2） 知识图谱构建

知识图谱在逻辑结构上主要分为：数据层和模式层。数据层包含大量的事实

（fact）信息，即（实体，关系，实体）或者（实体，属性，属性值）等三元组表示形式，将这些数据存储在图数据库（比如：开源的 Neo4j、Twitter 的 FlockDB、sones 的 GraphDB、中科天玑自主研发的 Golaxy Graph 等）中会构成大规模的实体关系网络，进而形成知识图谱。模式层是知识图谱的核心，建立在数据层之上，存储的是提炼后的知识。通常采用本体库来管理模式层，即使用本体库对公理、规则和约束条件的支持能力来规范实体、关系以及实体的类型和属性等对象之间的联系。本体库在知识图谱中的地位相当于知识库的模具，拥有本体库的知识图谱冗余知识较少[40]。

根据上述的知识图谱分层结构，知识图谱的构建方法主要有两种：一种是自底向上的构建方法；一种是自顶向下的构建方法。自底向上的构建方法是从开放链接的数据源中，采用自动或半自动的技术方法，抽取实体、属性和关系等信息，加入到知识图谱的数据层；然后将这些知识要素进行归纳组织，逐步往上抽象为概念，最后形成模式层。自顶向下的方法正好相反，首先从高质量数据（比如百科类网站）中抽取本体和模式信息，构建模式层；然后根据模式层的规则，从开放的数据集中抽取实体、属性和关系等信息，加入到数据层。

在知识图谱技术发展初期，大部分知识图谱产品主要采用自顶向下的构建方式，比如 Cyc、Freebase、DBpedia、YAGO 等。随着自动化知识抽取技术的不断成熟，以及公开采集的数据规模的不断扩大，现在的知识图谱大多采用自底向上的构建方式，比如谷歌的 Knowledge Vault、卡内基梅隆大学的 NELL 等。

知识图谱的构建是从互联网上结构化、半结构化和无结构化数据资源中，采用自动或半自动的技术抽取知识，并存入数据层和模式层的过程。自底向上的知识图谱构建是一个迭代更新的过程，涉及的技术主要包括：信息抽取、知识融合、知识加工。图 18 展示了知识图谱的体系架构[40]。

AMiner

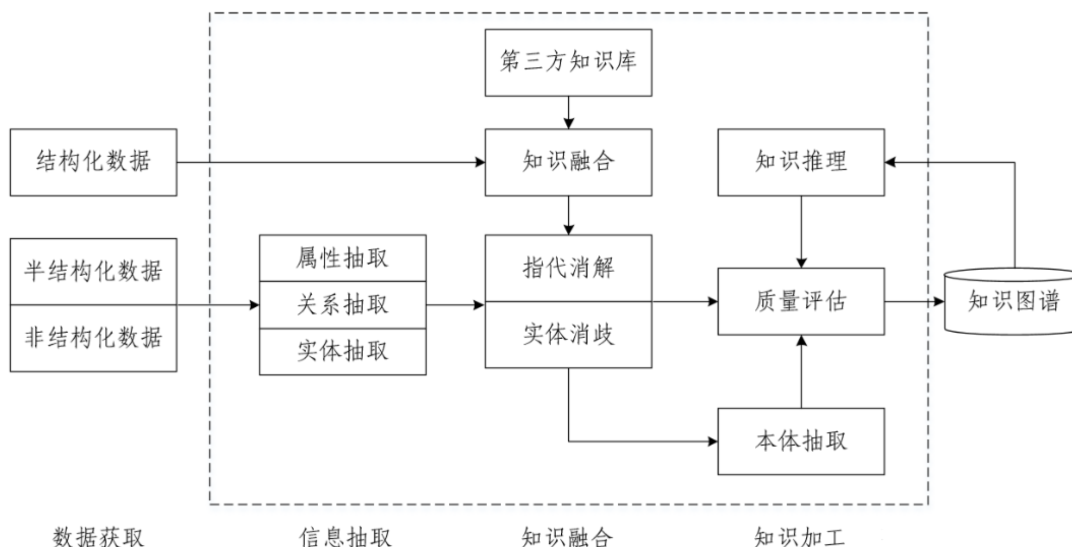


图 18 知识图谱的体系架构

- **信息抽取**

信息抽取（Information Extraction）是从自然语言文本中抽取实体、关系和实体属性等结构化信息的过程，涉及到的关键技术包括：实体抽取、关系抽取和属性抽取[41]。

✧ **实体抽取**

实体抽取（Entity Extraction），又称命名实体识别（Name Entity Recognition），是指从自然语言文本中自动识别出命名实体，目的是构建知识图谱中的“节点”。实体抽取的质量（准确率和召回率）对信息抽取后续任务的效率和质量影响极大，因此是信息抽取中最基础和关键的部分。传统的实体类别主要包括三大类和七小类：实体类（包括人名，地名，机构名）、时间类（日期，时间）、数字类（货币、百分比）。图 19 给出了实体抽取的样例，左侧图片中的下划线加粗的字符串展示了各类实体。在处理实际问题时，抽取的实体类别需要由特定的任务需求制定，比如医疗实体主要面向疾病、药物、症状、医疗设备等名称。

AMiner

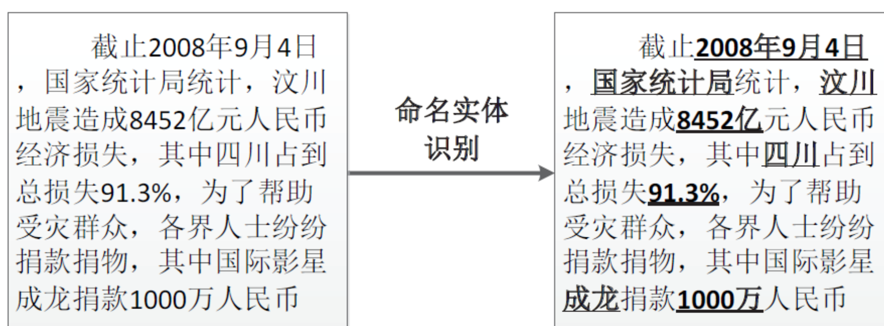


图 19 实体抽取样例

由于时间、货币、百分比、日期等命名实体具有规律性的构成成分，可以很容易识别，一般使用启发式规则、正则表达式等方法，研究人员对该类别实体研究较少。而人名、机构名和地名形式多变、用字灵活、内部结构复杂，识别难度较大。比如人名“杜甫”，具有多种名称表示，包括“杜工部、杜子美、李杜”等。对于机构名来说，经常使用简称，或者英文首字母来代替。比如“智谱·AI”的全称为“北京智谱华章科技有限公司”，英文简称为“Knowledge Atlas”。此外，文本中的实体指称项还和上下文密切关联，同一实体名称在不同语境下指代不同的事物，或者在某些语境下是实体，在其他语境下就不是实体。比如“河南”在地名介绍中表示一个省份，在地质介绍中可能表示河的南边。因此，研究者专注于人名、结构名、地名等三类命名实体的识别研究，并举办了多种国际会议来推动命名实体识别技术的发展，包括 MUC[42]、SigHAN[43]、CoNLL[44]、和 ACE[45]等。

AMiner

早期的实体抽取方法主要采用人工制定的启发式规则，比如 Rau 采用人工制定的启发式规则，从文本中自动抽取公司名称[46]。基于规则的方法不仅需要耗费大量的人力成本，而且可扩展性比较差，难以满足数据变化的需求。

为了克服基于规则方法的局限性，研究者们尝试使用统计机器学习方法来解决实体抽取问题。该方法利用实体的内部特征以及上下文特征，将实体抽取任务转化为多类分类或序列标注问题，并训练标注或分类模型。实体的序列标注主要包括 BIO 和 BIOES 两种。其中，B，即 Begin，表示开始；I，即 Intermediate，表示中间；E，即 End，表示结尾；S，即 Single，表示单个字符；O，即 Other，表示其他，用于标记无关字符。比如将“小明在北京大学的燕园看了中国男篮的一场比赛”这句话进行标注，结果就是：[B-PER，E-PER，O，B-ORG，I-ORG，

I-ORG，E-ORG，O，B-LOC，E-LOC，O，O，B-ORG，I-ORG，I-ORG，E-ORG，O，O，O，O，O]。

基于统计机器学习的实体抽取方法采用的学习模型主要包括条件随机场模型（Conditional Random Field, CRF）[47]、隐马尔科夫模型（Hidden Markov Models, HMMs）[48]、最大熵模型（Maximum Entropy Model, MEM）[49]、支持向量机（Support Vector Machine, SVM）[50]、鲁棒风险最小化模型（Robust Risk Minimization Model, RRMM）[51]、以及转化学习模型（Transformation-Based Learning Model, TBLM）[52]。统计机器学习模型依赖于训练语料，语料规模越大、覆盖面越广，模型学习能力就越高。但是由于各种客观原因，比如语料资源有限，在标注训练语料时，容易出现数据不平衡问题，即不同类别的语料数量相差很大，导致模型学习性能受限。而且，特征定义和语料标注一般采用手工方式，一方面增加了人工成本，另一方面定义的特征集合和标注的训练语料规模有限，降低了模型的学习效果。

随着信息技术的高速发展，网络上的用户生成文本呈现爆炸性增长，极大丰富了实体名称和类别的多样性。传统的面向单一领域的实体抽取方法由于实体类别限定的约束，无法满足实际应用的需求。学术界开始关注开放域的实体抽取问题，即自动从领域开放的互联网数据资源中抽取表示实体的所有文本指称项。该任务的主要挑战是实体类别无法预先定义，并且处理的语料资源的规范性无法保障。Lu 等人提出了一种超图模型（Hypergraph Model），通过将实体指称项的每一次出现都使用超图表示，并采用一种对数线性（log-linear）方法来同时进行指称项抽取和分类，有助于识别指称项的多层嵌套形式[53]。Radu 等人将实体指称项识别任务转化为多分类问题，并提出了多对一模型、级联模型和组合模型等三种分类模型，来执行实体指称项的检测和类别识别[54]。此外，研究者也将统计机器学习技术用于实体指称项识别任务中。Li 等人将实体指称项识别任务转化为序列数据的标注问题，并使用线性条件随机场模型来执行实体指称项的边界检测和类别识别[55]。Kadri 等人利用支持向量机模型中的一对多的分类特性，来执行中英文文本中实体指称项识别任务[56]。

✧ **关系抽取**

关系抽取（Relation Extraction）是指从自然语言文本中自动抽取实体对之间

AMiner

的关联关系。根据关系类别是否限定，关系抽取技术可以分为传统的关系抽取方法和开放域的关系抽取方法。传统的关系抽取方法大多采用统计机器学习技术，即根据预先定义的关系类别，标注训练语料，将关系抽取任务转化为多类分类问题。Kambhatla 使用多种词汇、句法和语义特征，训练一个最大熵分类器，来应对 49 种关系类别[57]。Zhou 等人基于 Kambhatla 的研究，探索了更多的特征，包括词汇、词组、语义资源等特征，并利用 SVM 分类器，来提高关系抽取的性能[58]。

为了降低人工标注训练语料的成本，以及解决训练语料数据不平衡问题，一些研究者使用自扩展、主动学习、标签传播等半监督方法来解决关系抽取问题。基于自扩展方法需要大规模的未标注语料，和每种关系类别下少量的种子关系实例。基于种子实例，自扩展算法会抽取相同类别的关系实例，利用学习的关系模式，来抽取更多的关系实例。DIPRE[59]是最早的基于自扩展算法的关系抽取方法，利用关系模式来学习关系元组，反过来，关系元组也可以用于学习关系模式。为了降低语料标注的成本，主动学习技术被机器学习领域广泛使用。主动学习的优势在于性能可以与监督方法相媲美，因而被一些研究者用于关系抽取任务上。Sun 和 Grishman[60]提出了 LGCo-Testing，基于主动学习方法 Co-testing，使用局部特征训练最大熵分类器，以及使用全局特征训练最近邻分类器。标签传播算法是一种基于图的半监督方法，将标注和未标注的实例表示为结点，实例之间的相似度表示为边，通过带权边，结点的标签信息会被传播到邻近结点，在传播过程中，通过推理来对未标注数据的进行标签划分[61]。Chen 等人最早将标签传播算法应用到关系抽取任务，将关系实例表示为结点，并分配一个特征向量，特征向量之间的相似度作为边[62]。

传统的关系抽取方法需要人工预先设置关系类别，导致关系类别有限。互联网技术的快速发展，引发了用户生成文本大量激增，显著提高了关系类别的丰富性，会出现一些未知的、随着时间演化的关系类别。此外，需要人工参与的方法，比如标注训练数据、制定模式、选择特征等，通常是任务依赖的，无法较好地应用到其他领域。为了克服以上限制，研究者提出了开放域的关系抽取（Open Relation Extraction）技术，从海量的不限特定领域的网络文本中自动抽取实体对，

AMiner

以及描述它们之间语义关系的指示词，并以（实体 1，关系指示词，实体 2）等关系三元组格式输出。比如，给定文本“杨幂携员工热巴出席芭莎晚会”，抽取的关系三元组为（杨幂，员工，热巴）、（杨幂，出席，芭莎晚会）、（热巴，出席，芭莎晚会）。

TextRunner 是第一个开放域关系抽取系统，由 Banko 等人于 2007 年提出，基于词性和名词性短语等特征训练了一个朴素贝叶斯分类器来抽取实体对之间的关系[63]。后续的工作证明使用分类器对建模文本中的序列信息很有成效，比如条件随机场[64]，马尔科夫逻辑网络[65]。Wu 等人提出的 WOE 系统采用了维基百科作为训练语料，以及依存关系特征等句法特征构建分类器。实验表明，使用依存关系特征可以产生比 TextRunner 更好的性能[66]。ReVerb[67]首先制定句法和词典约束来识别以动词为中心的关系词语，然后将输入语句分割为（论元 1，关系词，论元 2）三元组格式。该方法只利用了词性标注和正则表达式匹配，适用于大规模语料，不受语言限制。对于多语言的开放式关系抽取任务，Gamallo 等人提出了一种基于规则的依存关系解析器来抽取英文、西班牙语、葡萄牙语、加利西亚语等语料中的关系。对于每个句子，他们将其分割为动词性子句，并识别与该动词相关的成分，以及它们的功能，包括主语、直接宾语、定语、介词补语，并基于子句构成，设计一组规则来抽取关系元组[68]。

✧ **属性抽取**

属性是描述实体内在本质必不可少的信息，比如人物的年龄、身高，商品的生产日期、构成成分，机构的创立时间、创办者等，对知识图谱构建、用户意图理解、问答系统、推荐系统等多种智能应用具有重要的价值，引起了大量研究者的关注。2009 年，TAC 会议发布的知识库填充任务（Knowledge Base Population）中就包含实体属性值抽取任务，即槽填充（slot filling），至今已举办九届，并从英语扩展到汉语、西班牙语等多种语言。槽填充是指给定一组实体列表和属性名列表，从大量的无结构化文本数据中自动抽取实体属性的取值（即属性值[69]。由于属性值的构成形式和类别比较复杂，一些研究工作为了简化抽取，多会在限定条件下进行。比如面向医疗、电商等特定领域，针对人物等特定类别的实体，抽取实体型、数量型等特定类别的属性值。这些方法虽然会取得不错的抽取性能，

AMiner

但是可扩展性低下，无法满足实际的应用需求。

根据人工参与的程度，现有的实体属性值抽取方法大致可分为三种。第一种是以模式匹配为主的无监督方法，第二种是以训练学习模型为主的监督方法，第三种是通过自动构建训练语料而产生的弱监督方法。

基于模式匹配的属性抽取方法中，词汇—句法模式和基于依存路径的模式被较多研究者使用。其中，词汇—句法模式[70][71][72]是根据实体和属性信息的分布规律，直接从描述语句中获取。比如，“height of Yao Ming is 229 cm”（姚明的身高是 229 厘米）的句法模式是“attribute of entity is value”（实体的属性是实值），目前应用比较广泛。基于依存路径的模式[73]是将描述实体及其属性的语句进行依存关系解析，从中选择实体到属性名的依存路径作为模式，该模式是一个由词语和依存关系标签组合的字符串。比如，“John （entity） is very glad to marry Lily （attribute value）”（乔治（实体）很高兴和丽丽结婚（属性值））的依存模式为“nsubj_R <glad[JJ]> xcomp_L<marry [VB]> dobj_L <entity_first>”，其中“nsubj、xcomp、dobj”为使用斯坦福自然语言处理工具[74]解析的词语之间的依存关系标签，“_R, _L”表示依存关系的核心词是在右边还是左边，“<entity_first>”表示实体和属性名在句子中的顺序。由于句法模式对文本的规范性和规律性要求比较高，因此该方法主要面向英文语料，针对中文语料的研究匮乏。由于基于模式匹配的方法依赖于句法分析、精确匹配等技术，因此会导致数据稀疏问题。为了克服该问题，一些研究者提出借助搜索引擎，利用网络数据的冗余性来提高召回率，然而搜索引擎的查询限定制约了实际应用。

监督的抽取方法主要基于统计的机器学习和深度学习技术。Ye 等人首先利用从《知网》中获取的属性触发词实现对语句的分类，然后分别对每个属性训练一个 SVM 模型来识别属性值[75]。Huang 等人利用 SVM 分类器获取属性的描述语句，然后使用 CRF 模型从中识别属性值[76]。Liu 等人提出了一种基于注意力机制的循环神经网络模型，并将对齐信息（Alignment Information）融入到模型中，提高了属性值抽取性能[77]。Yao 等人使用 Word2vec 语言模型获取词向量表示，作为循环神经网络模型的输入，来从中文语料中进行人物属性抽取[78]。Adel 等人提出了一种语境感知的卷积神经网络模型，来对属性值抽取任务中的关系分类

AMiner

问题进行研究[79]。刘等人提出了一种结合全局特征和局部特征的感知机学习模型，用于抽取非结构化文本中命名实体的属性值[80]。

有监督的方法由于需要人工标注训练语料，因此在实际应用中受到很大局限性。为了降低数据标注成本，部分研究者将注意力转向弱监督方法研究上。Wu 等人利用维基百科信息框中的属性信息，从正文中回标语句，来自动构建训练语料，然后训练 CRF 模型抽取属性值[81]。Probst 等人将属性值抽取转化为分类问题，通过利用词语的互信息，制定启发式规则，来获取包含属性值和属性名的文本片段，并结合预构建的属性值词典自动标注数据，作为朴素贝叶斯分类器的训练语料，并提出了一种半监督的 co-EM 算法来提高分类效果[82]。He 等人借助百度百科和百度搜索自动获取训练语料，训练了一个双向长短期记忆网络模型，来从中文语料中进行实体属性抽取[83]。曾等人利用百度百科中的信息框对其正文进行回标产生训练数据，并提取多维度文本特征，训练 CRF 模型，来执行实体属性抽取任务[84]。

- **知识融合**

由于信息抽取任务面向多种数据源，而不同数据源对同一对象的描述信息存在多样性等问题，导致抽取的实体、关系和属性等存在冗余信息。为了保证知识图谱的质量，有必要对信息抽取的结果进行清理和整合，使用的关键技术包括实体消歧和共指消解。

✧ **实体消歧**

实体消歧（Entity Disambiguity）是指明确自然语言文本中的实体指称项指代的真实世界事物，主要针对实体名称的多样性和歧义性问题提出的。实体多样性是指同一实体在不同自然语言文本中具有多样的表达形式。比如，篮球运动员“科比”，又称为“小飞侠”、“黑曼巴”、“科神”等。实体多样性现象一方面丰富了实体的名称表示形式，另一方面由于不规律性的表示形式，增加了实体识别的难度。实体歧义性是指同一实体指称项在不同语境信息中，指代不同的真实世界实体。比如，人名“乔丹”，在篮球描述数据中，表示著名的 NBA 篮球运动员，而在人工智能相关的文档中，则表示机器学习专家。如果不结合语境信息，我们无法判

AMiner

别有歧义名称的真实指向。实体的多样性和歧义性问题给自然语言理解、信息检索和知识图谱构建与更新等任务带来了挑战。

目前，实体消歧技术可分为基于聚类技术的消歧方法和基于实体链接技术的消歧方法。基于聚类技术的实体消歧方法，其基本思路是通过度量实体指称项之间的名称或语境相似度，将指向同一实体的指称项聚为一类。传统的方法主要采用基于词袋模型[85][86]，利用同现的语境信息，或者抽取实体属性信息来度量实体指称项之间的相似度。然而这类方法忽略了大部分的语义知识，比如概念间的语义关联、实体间的社会化关联等。因此，后续研究主要针对如何挖掘准确全面的语义知识来表示实体指称项的问题寻找突破[87][88][89]。比如从维基百科、Freebase等外部知识源中挖掘隐藏的语义知识，或者集成来自不同知识源的语义知识。然而，知识源的多源异构性，以及大部分语义知识隐藏在复杂结构中等问题带来了研究困难。

由于基于聚类技术的实体消歧方法只能挖掘指向同一实体的多种名称指称项，而忽略了探索这些指称项背后的实体指代，即仅解决了实体的多样性问题，因而无法满足互联网实际应用的需求。针对以上不足，研究者提出了基于实体链接技术的消歧方法（简称实体链接）。实体链接是指，通过将自然语言文本里的实体指称项链接到给定知识库中的正确词条，来明确其真实含义，从而达到实体消歧的目标[90]。实体链接（Entity Linking）技术利用知识库中词条来明确实体指称项的真实指代，可以同时解决实体名称多样性和歧义性问题。图 20 展示了一个实体链接的例子[91]。通常地，一个常规的实体链接系统包含三个模块：候选实体产生、候选实体排序和空链接指称项预测[90]。下面，我们将根据各个模块介绍相关的研究方法。

1）候选实体产生。对于每个实体指称项，系统首先需要从给定知识库中寻找该指称项可能链接到的候选实体列表。候选实体列表的规模会影响实体链接的性能，规模大会提高包含目标链接实体的概率，但是会增加计算开销。因而，在候选实体产生阶段需要综

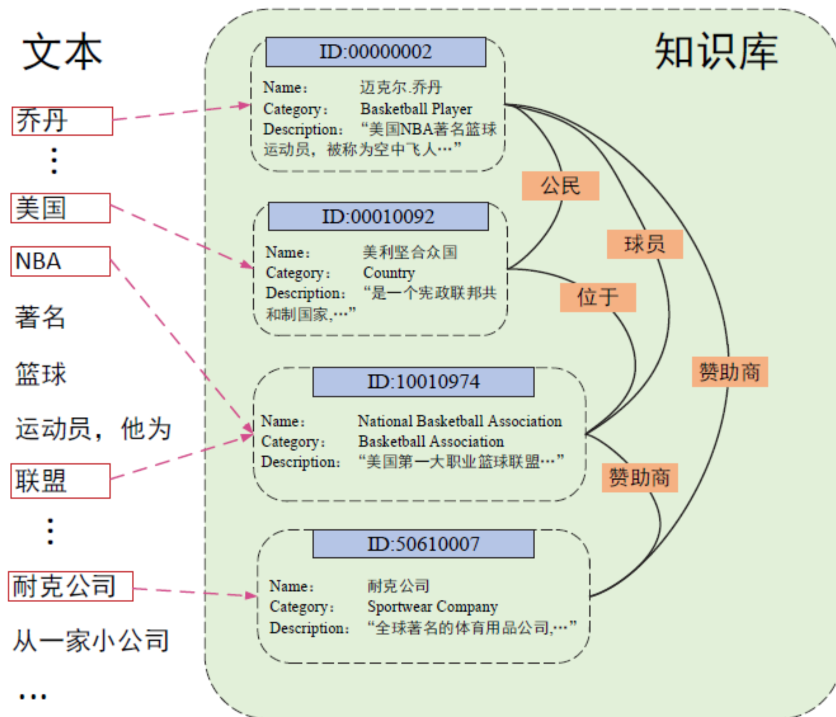


图 20 实体链接案例

合考虑计算准确性和成本开销问题。产生候选实体的方法主要基于度量实体指称项和知识库中实体之间的名称相似度。为了实现以上目标，基于同义词表构建方法、基于名称扩展方法、以及基于搜索引擎方法被相继提出。

基于同义词表构建方法是候选实体产生的主要方法，并被多种实体链接系统采用[92][93][94]。维基百科为产生候选实体提供了一组丰富的有效特征，包括实体页面、消歧页面、重定向页面、锚文本等。相关工作通过组合这些特征来构建一个关于实体名称不同表示形式的词典，包括别名、简称、昵称、拼写变形、易混淆名称等。然后使用字符串完全匹配、模糊匹配等算法，度量实体指称项与词典中名称的相似度，来获取候选实体。然而，该种方法容易受到名称拼写错误的影响，降低了实体指称项与同义词表中词条的名词相似度。

为了解决实体指称项的名称是全称的片段或缩写形式等问题，一些实体链接系统会从实体指称项的描述文档中挖掘其他的名称表示，比如全称。然后利用扩展的名称形式来产生候选实体[92][95]。基于搜索引擎的方法是指借助搜索引擎，利用网络信息来识别候选实体。比如，将实体指称项放入谷歌搜索引擎中，并将返回的维基百科页面词条作为候选实体[96][97]。

AMiner

2）候选实体排序。由于候选实体列表通常包含两个或多个候选实体，为了获得精准的链接目标，研究者探索多种方法对候选实体列表进行排序。根据同时排序实体指称项的个数，候选实体排序方法可以分为单一排序方法[87][98]、协同排序方法[99][100]和协作排序方法[101][102]。单一排序方法将一篇文档中提及的实体指称项进行单独处理，没有考虑它们之间的关联关系，通常度量实体指称项的语境信息和候选实体的描述文档之间的相似度，来实现候选实体排序。协同排序方法假设一篇文档由属于一个或多个主题的实体指称项组成，它们之间存在关联关系，利用主题一致性的思想，对一篇文档中提及的所有实体指称项实现同时链接。协作排序方法针对需要链接的实体指称项，从其他文档中寻找名称相似或语境相似的实体指称项，利用交叉文档来扩展被链接的实体指称项的语境信息，来协助排序候选实体。

根据使用的候选实体排序模型，候选实体排序方法可以分为：二分类方法、学习排序方法、概率模型方法、基于图方法。二分类排序方法[97][103][104]是指将候选实体排序任务转化为二分类问题。给定一对实体指称项和候选实体，通过训练分类器，如支持向量机（Support Vector Machine）、逻辑回归分类器（Logistic Regression Classifier）、朴素贝叶斯（Naive Bayes）等，来判断该候选实体是否为链接目标。学习排序方法融入了候选实体之间的关联关系，通过训练一个学习排序模型，来实现对候选实体的排序[96][105]。此外，一些研究者基于“一篇文档中的所有实体指称项描述同一主题”的思想，提出了概率模型方法，来同时链接文档中的所有实体指称项[106][107][108]。基于图的排序方法充分考虑了实体指称项的局部和全局特征，通过构建相关图来表示实体指称项和候选实体，实现对同一文档中的所有实体指称项协同链接[92][99][109]。

AMiner

3）空链接指称项预测。空链接指称项预测模块的目标是处理没有目标链接实体的实体指称项。一些研究工作通常使用该模块验证候选实体排序模块输出的链接目标是否正确。为了判别实体指称项是否为空，最简单的预测方法是看候选实体列表是否为空[110][111]。若为空，则该实体指称项为空链接，标记为 NIL。一些实体链接系统也通过设置空链接阈值来判定，计算实体指称项与其每个候选实体的相关度，若最高的相关度分数低于空链接阈值，则判定该实体指称项为

NIL[87][92]。大部分实体链接系统使用监督的机器学习方法来预测空链接实体指称项，包括二分类方法和学习排序方法。基于二分类方法预测实体指称项和其最相关候选实体的类别标签，若为负，则实体指称项为 NIL。基于学习排序方法在候选实体列表中增加一个空链接实体，如果排序模型将该空链接实体作为最后输出，则判定实体指称项为 NIL。

✧ **共指消解**

共指消解（Coreference Resolution）技术主要用于解决多个指称项对应同一实体对象的问题。在一段文本中，多个名词和代词指称项可能指向的是同一实体对象。比如，"马云被很多网友称为"马爸爸"，因为他资助了很多贫困家庭"，这句话中的实体指称项"马云、马爸爸、他"均指向同一人物。利用共指消解技术，可以将这些指称项关联（合并）到正确的实体对象，由于该问题在信息检索和自然语言处理等领域具有特殊的重要性，吸引了大量的研究。

共指消解需要考虑的特征主要分为以下几类：词汇、距离、一致性、语法、语义等。词汇特征主要考虑两个指称项字符串的匹配程度，一般而言字符串相同程度越高的指称项共指概率越大。距离特征主要考察两个指称项的句子距离，这个主要依据是共指事实上也是一种局部性的替代关系，越是临近的指称项之间共指概率越大。一般而言，两个指称项相隔超过三个句子，共指的可能性就会很小了。一致性特征详细可以分为性别、单复数、语义类别等是否一致。这组特征主要起到筛选的作用。语法关系用来判断两个指称项的语法角色之间的关系，由于对句子深层的语法分析还很难办到，这里主要采用的是一些基于特定模板的方法，例如判断两个指称项之间是否被逗号格开或者相邻等来决定是否具有同位关系。语义特征主要是考察两个指称项在语义类别不一致时是否满足上下位或者同义、近义关系。这种判断主要依赖于具体的语言学词典，例如英文上的 WordNet、中文上的 HowNet 等。

基于自然语言处理的共指消解是以句法分析为基础的，代表方法是 Hobbs 算法和向心理论（Centering Theory）。Hobbs 算法是最早的代词消解算法之一，主要思路是基于句法分析树进行搜索，因此适用于实体与代词出现在同一句子中的场景，有一定的局限性[112]。向心理论的基本思想是：将表达模式（Utterance）视

AMiner

为语篇（Discourse）的基本组成单元，通过识别表达式中的实体，可以获得当前和后续语篇中的关注中心（实体），根据语义的局部连贯性和显著性，就可以在语篇中跟踪受关注的实体[113]。

随着统计机器学习方法被引入该领域，共指消解技术进入了快速发展阶段，McCarthy 等人首次将 C4.5 决策树算法用于解决共指消解问题[114]。Bean 等人在共指消解算法中融入了语义背景知识，首先利用 Utah 大学发布的 AutoSlog 系统从原始语料中抽取实体上下文模式信息；然后应用 Dempster—Shafer 概率模型对实体模式进行建模[115]。为了解决基于统计机器学习的共指消解方法中训练数据的（特征）稀疏性和难以在不同的概念上下文中建立实体关联等问题，Pantel 等人基于 Harris 提出的分布相似性模型，提出了一个新的实体相似性测度模型，称为术语相似度（Term Similarity），借助该模型可以从全局语料中得到所有术语间的统计意义上的相似性，据此可以完成实体合并，达到共指消解的目的[116]。

- **知识加工**

AMiner

通过前述的信息抽取技术从互联网上的数据资源中抽取的实体、关系和属性等信息，经过知识融合技术可以消除实体指称项和世界真实事物之间的歧义，得到实体关系三元组等事实表达形式。但是如果得到具有逻辑层次和高质量的知识图谱，还需要经历知识加工的过程，涉及的技术包括：本体构建和质量评估。

✧ **本体构建**

本体构建主要针对知识图谱的模式层，通过对公理、规则和约束条件的支持能力来规范实体、关系以及实体的类型和属性等对象之间的联系。本体库在知识图谱中的地位相当于模具，拥有本体库的知识图谱冗余知识较少，结构层次强。

本体的五个基本构成元素包括：类（Class）或概念（Conception）、关系（Relations）、函数（Functions）、公理（Axioms）、实例（Instance）。

1）类（Class）或概念（Conception）

类，是对象的集合；概念是广义的概念，除了可以是一般意义上的概念以外，也可以是任务、功能、行为、策略、推理过程等，其定义一般包括概念的名称，

以及对该概念的自然语言描述。本体中的这些概念通常构成一个继承的分类层次。

2）关系（Relations）

关系是领域中概念之间的交互作用，形式上定义为 n 维笛卡儿积的子集：R：C1×C2×…×Cn。概念间的关系多种多样，但是那些具有普遍性的关系更值得关注，例如表达概念间部分与整体的关系 Part-of、表达概念间的继承关系 Kind-of、表达概念的实例与概念间的关系 Instance-of、表达某个概念是另一个概念的属性 Attribute-of。

3）函数（Functions）

函数是一种特殊的关系，关系的前 n－1 个元素可以惟一确定第 n 个元素。一般地，函数采用 F：C1×C2×…×Cn-1→Cn 的形式表示，如 MotherOf 就是一个函数，MotherOf（x,y）表示 y 是 x 的母亲，即 x 确定 y。

4）公理（Axioms）

公理是一些永真式的描述。更具体地说，公理是领域中在任何条件下都成立的断言，如概念乙属于概念甲的范围。

5）实例（Instances）

实例代表属于某概念/类的基本元素，即某概念/类所指的具体实体。从语义上分析，实例表示的就是对象，而概念表示的则是对象的集合，关系对应于对象元组的集合。概念的定义一般采用框架（Frame）结构，包括概念的名称、与其他概念之间的关系的集合以及用自然语言对该概念的描述。

在实际建模过程中，不一定要严格地按照上述 5 类基本建模元语来创建本体，概念之间的关系不限于上面列出的 4 种基本关系，可以根据领域的具体情况定义相应的关系，以满足应用的需要。本体构建的技术主要分为两种：一种是人工构建，然而这种方式需要消耗大量的人工成本，而且针对特定领域，难以找到合适的专家。另一种是数据驱动的自动化本体构建，包含 3 个阶段：实体并列关系相似度计算、实体上下位关系抽取、本体生成[40]。

实体并列关系是指任意给定的两个实体是否属于同一语义类别，比如“中国”

AMiner

和“美国”都属于国家类别，而“中国”和“北京”不属于同一类别。主流的实体并列关系相似度计算方法主要包括模式匹配法和分布相似度法。模式匹配法采用预先定义实体对模式的方式，通过模式匹配取得给定关键字组合在同一语料单位中共同出现的频率，据此计算实体对之间的相似度。分布相似度是利用实体的上下文信息将实体表示为向量形式，通过度量两个实体向量之间的相似度，来判别两个实体是否属于同一语义类别。

实体上下位关系是确定概念之间的隶属关系，比如词组（中国，国家）构成上下位关系，中国是下位词，国家是上位词。主要研究方法是基于语法模式来抽取实体对，然后利用百科类网站提供的概念分类知识训练概率模型来判定隶属关系和区分上下位词。

本体生成的主要研究是实体聚类方法，利用主题知识对实体进行层次聚类来得到本体结构。然而由于信息抽取的实体信息缺乏上下文语境，导致很多统计模型不可用。为了解决该问题，Wang 等人提出了一个基于词共现网络的主题聚类和上位词抽取模型，实现了基于短文本的主题聚类[117]。Liu 等人则采用贝叶斯模型对实体关键词进行分层聚类，经过改进的算法具有近似线性的复杂度 0（nlog（n）），能够在 1h 内从 100 万关键词中抽取出特定领域的本体[118]。

✧ **质量评估**

由于数据资源的质量难以保证，以及受到技术水平的限制，从互联网中抽取的实体、关系和属性等信息可能存在错误，因此在加入知识图谱之前，需要对它们的可信度进行评估，舍弃置信度低的知识，以保证知识图谱的质量。现有的知识图谱质量评估方法主要是利用数据来源的可信度来判别知识的质量。谷歌的 Knowledge Vault 项目从全网范围内抽取结构化的数据信息，并根据某一数据信息在整个抽取过程中出现的次数对其可信度进行评分，然后利用从可信知识库 Freebase 中得到的先验知识对先前的可信信息进行修正，实验结果表明，这一方法可以有效降低对数据信息正误判断的不确定性，提高知识图谱中知识的质量[20]。谷歌提出了一种依据用户的贡献历史和领域，以及问题的难易程度进行自动评估用户贡献知识质量的方法。用户提交知识后，该方法可以立刻计算出知识的可信度，使用该方法对大规模的用户贡献知识的评估准确率达到了 91%，召回

AMiner

率达到了 80%[119]。除了利用数据来源的可信度，Fader 等人在对 REVERB 系统的信息抽取质量进行评估时，采用人工标注方式对 1000 个句子中的实体关系三元组进行了标注，并以此作为训练集，得到了一个逻辑回归模型，用于对 REVERB 系统的信息抽取结果计算置信度[120]。随着数据规模的不断扩大，数据间的冲突日益增多，质量评估是知识图谱构建技术的重要组成部分，对知识质量评估问题的研究需要重点关注建立完善的质量评估技术标准和指标体系。

（3） 知识图谱存储

基于文献[121]，本节对知识图谱主流的数据模型、查询语言和存储机制分别展开介绍。

- **知识图谱数据模型**

本节重点介绍两个主流的知识图谱数据模型：RDF 图模型和属性图模型。通过以下介绍可以发现，RDF 图模型中一条三元组的谓语可以在另一条三元组中作主语或宾语，具有超图本质，而属性图中顶点和边属性不能再定义属性，因此 RDF 图模型的表达能力强于属性图模型。属性图模型虽然在理论基础方面不如 RDF 图模型完善，但是在工业界的应用范围更广阔，尤其是随着 Neo4j 等图数据库的应用，其获得了较强的用户认可度。

✧ **RDF 图模型**

RDF（Resource Description Framework，资源描述框架）[122],是 W3C 万维网联盟针对语义 Web（Semantic Web）制定的表示和交换机器可理解信息的标准数据模型[123]。RDF 使用 Web 标识符（URIs）来标识资源，RDF 使用属性和属性值来描述资源。其中，资源是可拥有 URI 的任何事物，比如 "http://www.w3school.com.cn/rdf"；属性是拥有名称的资源，比如 "author" 或 "homepage"；属性值是某个属性的值，比如 "David" 或 "http://www.w3school.com.cn"（请注意一个属性值可以是另外一个资源）。

在 RDF 图中,每个资源具有一个（HTTP URI）作为其唯一 id。资源、属性和属性值的组合可形成一个陈述（被称为陈述的主体、谓语和宾语）。RDF 图定义为三元组（s，p，o）的有限集合；每个三元组表示是一个事实陈述句（比如：

AMiner

"The homepage of is http://www.w3school.com.cn"），其中，s 是主语（比如 http://www.w3school.com.cn/rdf），p 是谓语（比如 homepage），o 是宾语（比如 http://www.w3school.com.cn）；（s，p，o）表示 s 与 o 之间具有联系 p，或表示 s 具有属性 p 且其取值为 o。

在图 21 所示的 RDF 图示例中，用椭圆表示资源，用矩形表示字面量；一条有向边及其连接的两个顶点对应于一条三元组，尾顶点是主语，边标签是谓语，头顶点是宾语。资源所属的类型由 RDF 内置谓语 rdf:type 指定，如三元组（James_Cameron,rdf:type,Director）表示 James_Cameron 是导演。

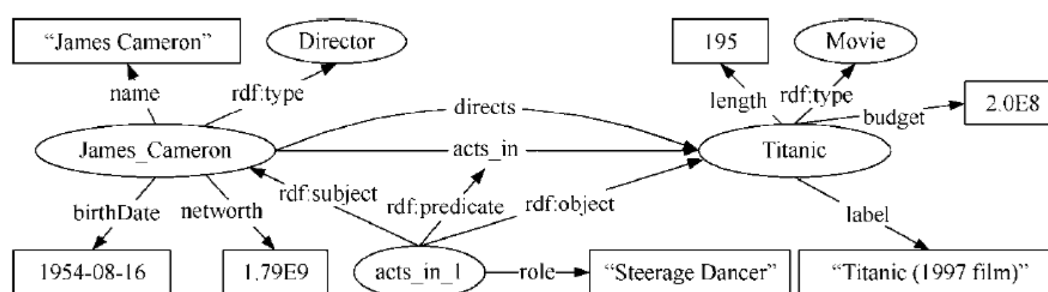


图 21 RDF图示例：电影知识图谱

AMiner

RDF 图是一种特殊的有向标签图，与普通有向标签图相比，RDF 图的特殊性在于，其三元组集合的本质使得一个三元组中的谓语也可作为另一个三元组的主语或宾语。反映在有向标签图中，即边亦可作为顶点，顶点与边交集非空。

✧ 属性图模型

与 RDF 图模型相比，属性图模型对于顶点属性和边属性具备内置的支持。目前，属性图模型被图数据库业界广泛采用，包括著名的图数据库 Neo4j。

属性图 G 可以形式化表示为 5 元组$(V, E, \rho, \lambda, \sigma)$。其中，（1）V 是顶点的有限集合；（2）E 是边的有限集合且$V \cap E = \emptyset$；（3）函数$\rho: E \rightarrow$（$V \times V$）将边关联到顶点对，如$\rho(e) =$（v_1, v_2）表示 e 是从顶点 v1 到顶点 v2 的有向边；（4）设 Lab 是标签集合，函数$\lambda:$（$V \cup E$）$\rightarrow Lab$为顶点或边赋予标签，如$v \in V$（或 $e \in E$）且 $\lambda(v) = l$（或 $\lambda(e) = l$），则 l 为顶点 v（或边 e）的标签；（5）设

Prop 是属性集合，Val 是值集合，函数σ（$V \cup E$）$\times$ Prop $\rightarrow$ Val为顶点或边关联属性，如$v \in V$（或 $e \in E$）、$\rho \in$ Prop且$\sigma(v,p) = val$（或 $\sigma(e,p) = val$），则顶点 v（或边 e）上属性 p 的值为 val。

如图 22 所示，属性图中每个顶点和边都具有唯一 id（如顶点 v1、边 e2）；顶点和边均可具有标签（如顶点 v1 上的标签 Director、边 e2 上的标签 acts_in），其作用基本相当于 RDF 图中的资源类型；顶点和边上均可具有一组属性，每个属性由属性名和属性值组成（如顶点 v1 上的属性 name="James Cameron"、边 e2 上的属性 role="Steerage Dancer"）。可以看出，利用边属性的定义，属性图示例增加了出演的角色信息，同时又没有改变属性图的整体结构（RDF 图中增加了顶点和边，改变了图结构）。

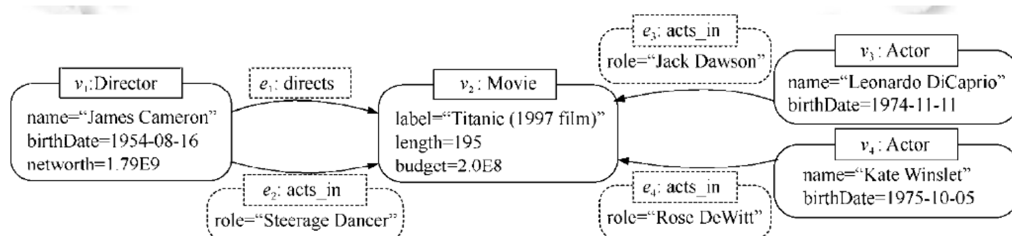


图 22 属性图示例：电影知识图谱

- **知识图谱查询语言**

针对上述的 RDF 图模型和属性图模型，本节分别介绍其代表性的查询语言 SPARQL 和 Cypher。

✧ **SPARQL**

SPARQL[124]的全称是 SPARQL Protocol and RDF Query Language，专门用于访问和操作 RDF 数据，是语义网的核心技术之一。从 SPARQL 的全称我们可以知道，其由两个部分组成：协议和查询语言。协议是指我们可以通过 HTTP 协议在客户端和 SPARQL 服务器（SPARQL endpoint）之间传输查询和结果，这也是和其他查询语言最大的区别。查询语言是指利用 SPARQL 语言来查询 RDF 数据。

SPARQL 查询是基于图匹配的思想。SPARQL 查询分为三个步骤：

1）构建查询图模式，表现形式就是带有变量的 RDF。

2）匹配，匹配到符合指定图模式的子图。

3）绑定，将结果绑定到查询图模式对应的变量上。

SPARQL 的常用关键词：

- ✓ SELECT：指定我们要查询的变量。在这里我们查询所有的变量，用*代替。
- ✓ WHERE：指定我们要查询的图模式。含义上和 SQL 的 WHERE 没有区别。
- ✓ FROM：指定查询的 RDF 数据集。我们这里只有一个图，因此省去了 FROM 关键词。
- ✓ PREFIX：用于 IRI 的缩写。
- ✓ FILTER：对变量取值进行约束。

比如：针对“巩俐参演的评分大于 7 的电影有哪些”语句的 SPARQL 的查询语句是：

```
SELECT ?n WHERE {
  ?s rdf:type :Person.
  ?s :personName '巩俐'.
  ?s :hasActedIn ?o.
  ?o :movieTitle ?n.
  ?o :movieRating ?r.
FILTER （?r >= 7）
}
```

AMiner

✧ Cypher

Cypher[125]是一个描述性的图形查询语言，允许不必编写图形结构的遍历代码对图形存储进行有表现力和效率的查询。Cypher 最初是图数据库 Neo4j 中实现的属性图数据查询语言，2015 年由 Neo4j 公司发起的开源项目 openCypher 进行了标准化，现在已应用到 Neo4j、SAP HANA Graph、Memgraph、Redis Graph

等图数据库产品中。

Cypher 的一个主要特点是使用“ASCII 艺术（ASCII art）”语法表达图模式匹配，是一种声明性模式匹配语言，遵循 SQL 语法，关注点是如何从图中找回结果，而不是怎么去做。Cypher 的代表性关键词如下：

- ✓ CREATE：创建节点、关系和属性。
- ✓ MATCH：检索有关节点、关系和属性数据。
- ✓ WHERE：过滤条件。
- ✓ DELETE：删除节点和关系。
- ✓ REMOVE：删除节点和关系的属性。
- ✓ ORDER BY：排序检索数据。
- ✓ SET：添加或更新标签。
- ✓ RETURN：返回查询结果。

比如，针对语句“获取 John 的朋友的朋友”，Cypher 查询语句如下：

```
CREATE john=node:node_auto_index（name = 'John'）
MATCH john-[:friend]->（）-[:friend]->fof
RETURN john, fof
```

- **知识图谱存储机制**

本节介绍两种典型的知识图谱存储机制：基于关系的知识图谱存储机制、原生知识图谱数据库的底层存储机制。

✧ **基于关系的知识图谱存储管理**

基于关系数据库的存储方案是目前知识图谱数据使用较多的一种存储方法。本节主要介绍三种具有代表性的基于关系的知识图谱存储方案，包括：三元组表、属性表和 DB2RDF。

1) 三元组表

AMiner

三元组表（triple table）是在关系数据库中建立一张具有 3 列的表，表模式是 triple_table（subject, predicate, object）。其中，subject、predicate 和 object 分别表示主语、谓语和宾语。三元组表是将知识图谱中的每条事实三元组存储为 triple_table 中的一行记录，如表 3 所示。

表 3 三元组表示例

subject	predicate	object
Titanic	length	195
Titanic	type	Movie
James_Cameron	birthDate	1954-08-16

三元组表是将知识图谱存储到关系数据库的最简单、最直接的方法，但是三元组表的行数与知识图谱的边数相等，从而导致将知识图谱查询翻译为 SQL 查询后会产生三元组表的大量自连接操作。当三元组规模较大时，多个自连接操作会影响 SQL 查询性能。

AMiner

2) 属性表

属性表（Property Table）的存储方案是将知识图谱中同种类别的主语及其谓语和宾语存到一个表中。属性表的行数是知识图谱中同种类别的主语数量，列数是不同谓语的数量。

针对 RDF 图示例中的内容，属性表的存储方案是分为 director 和 movie 等 2 个表，存储示例如图 23 所示。

director

subject	type	name	birthDate	networth
James_Cameron	Director	James Cameron	1954-08-16	1.79E9

movie

subject	type	label	budget	length
Titanic	Movie	Titanic(1997 film)	2.0E8	195

图 23 属性表存储方案示例

属性表存储方案的优点是：（a）克服了三元组表的自连接问题；（b）解决了水平表中列数目和空值数量过多的问题。缺点是：（a）规模较大的知识图谱中主

语的类别可能有成千上万个，需要建立对应的数据表，容易超过关系数据库的限制；（b）不同主语的谓语差异较大，容易导致空值问题；（c）属性表的一行一列只能存储一个值，无法应对知识图谱中同一主语和谓语具有多个宾语的情况。

3) DB2RDF

DB2RDF是IBM在2013年提出的一种面向实体的RDF知识图谱存储方案。该方案具有“列维度”灵活的优势，即将表的列作为谓语和宾语的存储位置,插入数据时,将谓语动态映射存储到某列而不将列与谓语进行绑定。

DB2RDF存储方案由dph表、rph表、ds表和rs表等4张表组成。dph（direct primary hash）是存储方案的主表,该表中一行存储一个主语及其全部谓语和宾语。如果一个主语的谓语数量大于 k（关系数据库支持的表中最大列数目），则一行不足以容纳一个主语实体，会出现溢出（spill）。spill列是溢出标志，spill列为0表示一行可以存储一个主语实体，spill 列为 1 表示一行不足以存储一个主语实体，需要在下一行存储该主语实体的其他谓语及宾语信息。如果一个谓语存在多个宾语，则需要引入ds（direct secondary hash）表，并在dph表的对应宾语处生成一个id值，将该id值和每个对应的宾语存储为ds表的一行。

dph表和ds表实际上存储了实体顶点（主语）的出边信息（从主语经谓语到宾语）。为了提高查询处理效率,还需要存储实体顶点的入边信息（从宾语经谓语到主语）。为此，DB2RDF方案提供了rph（reverse primary hash）表和rs（reverse secondary hash）表，它们的存储机制和dph表、ds表相同，这里不再赘述。

由于dph表和rph表支持将不同实体的相同谓语分配到相同列上，以及同一列中可以存储多个不同的谓语，因此 DB2RDF 方案具备“列共享”机制，在关系表中最大列数目上限的情况下可以存储远超出该上限的谓语数目,也能够有效地解决水平表方案中存在的谓语稀疏性空值问题。

✧ **原生知识图谱存储管理**

原生知识图谱存储是指专门为知识图谱而设计的底层存储管理方案。本节选择两个具有代表性的原生知识图谱存储管理方案进行介绍：一种是面向属性图的

AMiner

Neo4j 存储；另一种是面向 RDF 图的 gStore 存储。

1) Neo4j

Neo4j[126]是一个高性能的图形数据库，将结构化数据存储在网络而不是表中。Neo4j 图数据库遵循属性图模型来存储和管理其数据，采用 Cypher 查询语言。Neo4j 最大特色是具有"无索引邻接（index-free adjacency）"特性。所谓"无索引邻接"是指，每个顶点维护着指向其邻接顶点的直接引用，相当于每个顶点都可看作是其邻接顶点的一个"局部索引"，用其查找邻接顶点比使用"全局索引"节省大量时间。这就意味着图导航操作代价与图大小无关，仅与图的遍历范围成正比。

为了实现"无索引邻接"，Neo4j 将边放到核心位置，并将属性图中的顶点、边、标签和属性分开存储在不同文件中。正是这种将图结构与图上标签和属性分开存储的策略，使得 Neo4j 具有高效率的图遍历能力。图 24 给出了 Neo4j 中顶点和边记录的物理存储结构，其中每个顶点记录占用 15 字节，每个边记录占用 34 字节。

顶点记录中，inUse（第 0 字节）是记录使用状态标志的字节，表示该记录是正在使用中，还是已经删除并可回收用于装载新记录；nextRelId（处于第 1 字节—第 4 字节）是与顶点相连的第 1 条边的 id；nextPropId（第 5 字节—第 8 字节）是顶点的第 1 个属性的 id；labels（第 9 字节—第 13 字节）是指向顶点标签存储的指针，若标签较少会直接存储在此处；extra（第 14 字节）用于存储一些内部使用的标志信息。

边记录中，inUse（第 0 字节）的含义与顶点记录相同，表示是否正被数据库使用的标志；firstNode（第 1 字节—第 4 字节）和 secondNode（第 5 字节—第 8 字节）分别是该边的起始顶点 id 和终止顶点 id；relType（第 9 字节—第 12 字节）是指向该边的关系类型指针；firstPrevRelId（第 13 字节—第 16 字节）和 firstNextRelId（第 17 字节—第 20 字节）分别为指向起始顶点的前一个和后一个边记录指针；secPrevRelId（第 21 字节—第 24 字节）和 secNextRelId（第 25 字节—第 28 字节）分别为指向终止顶点的前一个和后一个边记录指针；以上指向前后边记录的 4 个指针形成了两个"关系双向链"；nextPropId（第 29 字节—第 32

AMiner

字节）是边的第 1 个属性 id；firstInChainMarker（第 33 字节）表示该边记录是否为“关系链”中第 1 条记录的标志。

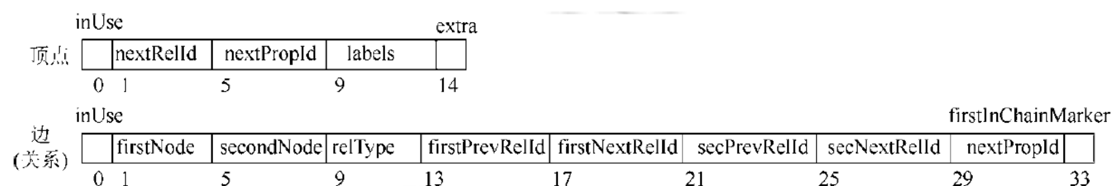


图 24 Neo4j 中顶点和边记录的物理存储结构

Neo4j 的特点如下：

- ✓ SQL 就像简单的查询语言 Neo4j CQL
- ✓ 它遵循属性图数据模型
- ✓ 它通过使用 Apache Lucence 支持索引
- ✓ 它支持 UNIQUE 约束
- ✓ 它包含一个用于执行 CQL 命令的 UI：Neo4j 数据浏览器
- ✓ 它支持完整的 ACID（原子性，一致性，隔离性和持久性）规则
- ✓ 它采用原生图形库与本地 GPE（图形处理引擎）
- ✓ 它支持查询的数据导出到 JSON 和 XLS 格式
- ✓ 它提供了 REST API，可以被任何编程语言（如 Java，Spring，Scala 等）访问
- ✓ 它提供了可以通过任何 UI MVC 框架（如 Node JS）访问的 Java 脚本
- ✓ 它支持两种 Java API：Cypher API 和 Native Java API 来开发 Java 应用程序

Neo4j 的优点为：

- ✓ 它很容易表示连接的数据
- ✓ 检索/遍历/导航更多的连接数据是非常容易和快速的
- ✓ 它非常易于表示半结构化数据
- ✓ Neo4j CQL 查询语言命令是人性化的可读格式，非常容易学习

AMiner

- ✓ 使用简单而强大的数据模型
- ✓ 它不需要复杂的连接来检索连接的/相关的数据，因为它很容易检索它的相邻节点或关系细节没有连接或索引

Neo4j 的缺点或限制为：

- ✓ AS 的 Neo4j 2.1.3 最新版本，它具有支持节点数，关系和属性的限制
- ✓ 它不支持 Sharding

2) gStore

gStore 是一个基于图的 RDF 三元组存储的数据管理系统，由北京大学、滑铁卢大学、香港科技大学基于 BSD 协议联合研发。gStore 图数据库的特点是：1）gStore 从图数据库角度存储和检索 RDF 知识图谱数据；2）gStore 支持 W3C 定义的 SPARQL1.1 标准，包含 Union、OPTIONAL、FILTER 和聚集函数查询，并支持有效的增删改操作；3）gStore 单机可以支持十亿三元组规模的 RDF 知识图谱的数据管理任务。

AMiner

gStore 主要技术手段包括：统一的结构和内容编码方法、一种高度平衡树 VS-tree 索引、基于索引的多级过滤机制。编码方式是首先将 RDF 数据图中每个资源的所有属性和属性值映射到一个二进制位串上，具体而言，gStore 对每个属性或属性值都定义了一个固定长度的位串，并将位串中所有位置设为 0；然后针对属性或属性值，按照标识符，利用若干个预先定义的字符串哈希函数将它们映射到若干个整数值（小于位串长度）,进而将位串上这些值所对应的位置修改为 1。

gStore 按照 RDF 图结构将所有位串组织成一棵签章树（signature tree），如果 RDF 图中两个实体之间有一条边，那么这两个实体在签章树上的对应点也连上一条边，且这条边被赋值为属性的编码。如此，gStore 中所有实体的编码就被组织成一种新的树形索引——VS*树（如图 25 所示）。VS*树被分为若干层,每一层都是 RDF 数据图的摘要。基于 VS*树，gStore 可以完成高效率的数据存储、更新与查询操作。当进行 SPARQL 查询处理时，将每个查询中的变量在这个 VS*树上进行检索，找到相应的候选解，然后再将这些候选解通过连接操作拼接起来[121]。

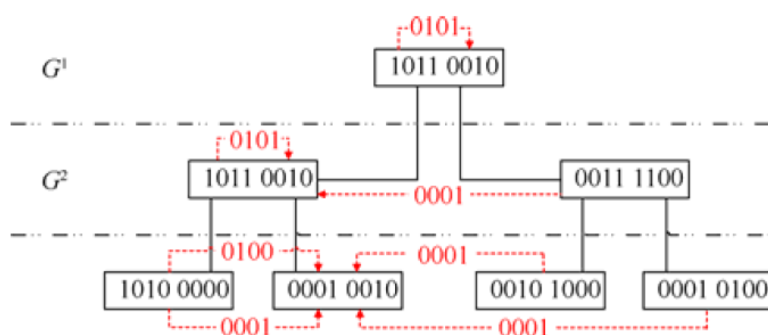


图 25 VS*树

2.1.4 知识图谱应用

Google 最初提出知识图谱是为了增强搜索结果，改善用户搜索体验，但知识图谱的应用远不止这些，按照应用方式可以分为语义搜索、知识问答、自然语言理解、推荐系统，以及基于知识的大数据分析与决策等。

■ 语义搜索：语义搜索是知识图谱最典型的应用，它首先将用户输入的问句进行解析，找出问句中的实体和关系，利用知识图谱定义良好的结构化大规模知识，挖掘实体和关系的深层含义，理解用户问句的意图，然后以有向图的方式提供满足用户需求的结构化语义内容，最后通过一定的形式将结果呈现到用户面前。语义搜索利用大规模知识图谱对搜索关键词和文档内容进行语义标注，改善搜索结果，如谷歌、百度和搜狗等公司在搜索结果中嵌入知识图谱，包括实体的结构化信息和相关实体的描述。

■ 智能问答：智能问答，是通过一问一答的形式，可以看作是语义搜索的延伸，通过对问句的语义分析，将非结构化问句解析成结构化的查询语句，并在知识图谱中匹配查询语句寻找答案。基于知识图谱的智能问答依赖于语义解析器的性能，在面对大规模、开放领域的域知识图谱时性能较差。近年来很多研究者开始研究更具鲁棒性的基于深度学习的知识图谱问答方法。

■ 推荐系统：借助知识图谱收集用户的兴趣偏好，产品的分类、属性、内容等，分析用户之间的社会关系，用户和产品的关联关系，利用推理算法，推断出用户的喜好和需求，从而为用户推荐感兴趣的产品或者内容。

■ 知识驱动的大数据分析与决策：利用知识图谱可以辅助行业和领域的大数据分析和决策，具体来说，利用知识图谱的知识，对知识进行分析处理，通过

AMiner

一定规则的逻辑推理，得出对于某种结论，为用户决断提供支持。美国 Netflix 公司利用基于其订阅用户的注册信息和观看行为构建的知识图谱，分析了解到 Fincher、Spacey 主演的作品表现都不错，以及英剧版的《纸牌屋》很受欢迎，因此决定拍摄了美剧《纸牌屋》，在美国及４０多个国家成为热门的在线剧集。

2.1.5 知识图谱研究问题与挑战

知识图谱追求的目标是花费更廉价的成本，使用更聪明的算法，获取大规模更高质量的数据，形成更好的判别或预测模型，实现认知层面的人工智能。纵观知识图谱的数据来源和相关技术研究现状，本节从数据资源、知识图谱表示、知识图谱构建、知识图谱存储等四个方面总结了未来知识图谱必须应对的挑战。

（1） 数据质量和规模。大数据时代，互联网每天产生海量的数据。抓取、下载、索引这些网页数据需要耗费大量的存储和网络资源，传统的数据挖掘和信息抽取算法在如此巨大的搜索计算空间中面临着效率瓶颈，难以应对大规模真实数据的动态变化。Web 数据中不仅蕴含丰富的价值信息，还混杂着大量的噪声数据，比如数据格式不一致、数据种类多样、命名模糊等，给信息抽取技术带来了很大挑战。

（2） 知识图谱表示。知识图谱是一种知识的图形化表示形式，需要为知识建立统一的语义空间，使得语义可计算，从而实现预测、推理以及推荐。虽然由 TransE 为代表的知识表示方法已经得到了应用和扩展，但仍然具有很大的挑战，包括需要设计融合更多本体特征的知识图谱表示学习算法，分析知识图谱表示学习与本体推理之间的等价性分析，以及进一步研究神经符号系统等。

（3） 知识图谱构建。构建知识图谱需要从不同种类的数据中抽取实体、关系和属性等知识，而数据种类不同，抽取知识的难度不一样，比如文本数据涉及到大量的自然语言理解的问题；图像数据涉及到大量视觉识别相关的问题；语音数据涉及到大量语音识别相关的问题。虽然结构化数据的知识抽取难度低，但是存在覆盖率的问题。而对于半结构化（比如网站）和非结构化数据（比如文本、图像、语音），不仅需要考虑包装器的生成和更新成本，还存在准确性的问题。

（4） 知识图谱存储。由于知识图谱的发展原因，其数据模型与查询语言的

AMiner

统一存在着一定难度。现有的知识图谱产品基于多种数据模型，比如 RDF 图、属性图。RDF 图模型虽然逻辑理论丰富，表达能力强于属性图模型，但是过强的理论信息影响了它在工业界的推广。而属性图在理论基础方面还不够完善，至今未形成一致公认的严格数学定义，虽然在多个图数据库中进行了应用，但尚未形成工业标准。此外，针对不同的数据模型，在不同的图数据库存在不同的查询语言，包括 SPARQL、Cypher、PGOL 等。数据模型与查询语言的不统一，不仅增加了数据库研发和维护的成本，还给用户学习带来了难度。

2.1.6 知识图谱未来研究方向

尽管人工智能依靠机器学习和深度学习取得了快速进展，但这些都是弱人工智能，要实现真正的类人智能，机器需要掌握大量的常识性知识，以人的思维模式和知识结构来进行语言理解、视觉场景解析和决策分析。如果未来的智能机器拥有一个大脑，知识图谱就是这个大脑中的知识库，能够打开人类的知识宝库，以结构化形式存储海量的人类知识，不仅对自然语言处理、信息检索和人工智能等领域产生深远影响，还为医疗、电力、金融等许多相关学科领域开启新的发展机会。

AMiner

虽然知识图谱在表示、构建和存储方面面临很多挑战，但是未来也有一些前景光明的研究方向。

（1） 已有工作对知识图谱中的关系划分为 1 对 1、1 对多、多对 1、多对多等 4 种类别，并针对不同的关系类别，研究了大量的知识表示模型。然而这些关系类别划分比较粗糙，难以有效解释知识的本质类型特点。而认知科学对人类知识类型的总结，有助于划分和处理知识图谱中的知识类型。因此，在知识图谱表示方面，未来有必要交叉研究人工智能和认知科学等领域，面向不同复杂关系类型的知识表示，有针对性的设计知识类型划分标准。

（2） 在知识图谱构建阶段，信息抽取是重要的一环。为了扩大知识图谱的规模和覆盖度，研究人员开始关注面向开放领域的信息抽取方法，虽然算法研究还在起步阶段，存在准确性和召回率低、扩展性不好等问题，但是从不限领域的数据资源中自动抽取知识是构建大规模知识图谱的必经之路，也是实现知识推理

等其他高层次技术的底层支撑。因此，开放域的信息抽取技术是知识图谱构建未来的研究重点。

（3） 在知识图谱存储方面，数据模型和查询语言尚不统一。统一的数据模型和查询语言，不仅减轻了数据库管理系统的研发成本，而且降低了用户设计、构建、管理和维护数据库的代价，同时降低了新用户的学习难度。因此，面向统一的知识图谱数据模型，研发统一的知识图谱查询语言，定义精确的语法和语义，是知识图谱存储未来的重要研究方向。

（4） 虽然一些神经网络在知识图谱表示和构建方面取得令人印象深刻的性能，但是它们在透明度和可解释性方面仍存在局限性。神经网络模型的不可解释性已经成为计算机领域顶级会议（如 NIPS）火药味十足的讨论话题。一些方法尝试将黑盒的神经网络模型和符号推理结合了起来，通过引入逻辑规则增加可解释性。此外，符号化的知识图谱具有形象直观的特性，为弥补神经网络在解释性方面的缺陷提供了可能。利用知识图谱解释深度学习和高层次决策模型，是当前值得研究的科学问题，可以为"可解释的 AI"提供全新视角的机遇。因为只有实现可解释性才可以说服人们相信预测结果，因此研究人员需要在可解释性和提升预测知识的可信度的方面做出更多的工作。

AMiner

2.2 认知推理

认知推理是实现认知智能不可或缺的部分，认知推理将结合人脑的推理过程，进一步解决复杂的阅读理解问题和少样本的知识图谱推理问题，协同结构化的推理过程和非结构化的语义理解。认知推理可以帮助机器获得多模感知能力，赋能海量任务。

知识图谱为认知推理提供了强有力的知识支撑，认知推理的底层实现是知识推理，而面向知识图谱的知识推理旨在基于已有的知识推理出新知识或者发现错误知识。因此，本节从基于知识图谱的知识推理角度出发，通过介绍知识图谱推理的概念、发展历程、关键技术、问题与挑战、未来研究方向等内容，为读者了解认知推理研究领域提供信息窗口。

2.2.1 知识图谱推理概念

AMiner

知识推理（Knowledge Reasoning, KR）作为人类求解问题的主要方法，一直以来备受关注。关于知识推理的基本概念，学术界给出了多种定义。文献[127]认为推理是从已知的事实出发，通过运用已掌握的知识，对各种事物进行分析、综合和决策，找出其中蕴含的事实，或归纳出新的事实的过程。文献[128]认为知识推理是指根据特定的规则和约束，从已存在的知识中获取新知识。文献[129]认为推理是为一系列能力的总称，包括有意识地理解事物的能力、建立和验证事实的能力、运用逻辑的能力以及基于新的或存在的知识改变或验证现有体系的能力。总体来说，知识推理是利用已有的知识，按照某种策略，获取新知识的过程。根据知识推理的相关定义，发现已有的知识是实现知识推理的前提，新知识是知识推理的结果。已有知识和推理得到的新知识的表示形式可以是句子（比如姚明的妻子是叶莉）、案例（比如小明以前参与了银行抢劫案件）、多元组表达形式（比如（姚明，夫妻，叶莉））等。

随着知识图谱的发展，对知识图谱的推理也越来越受到人们的关注。知识图谱推理可以形式化定义为：给定一个知识图谱$\mathrm{KG} =< \mathrm{E}, \mathrm{R}, \mathrm{T} >$和关系路径 P，E 和 T 表示实体集合，R 表示关系集合，R 中的边连接两个节点来形成一个关系三元组（$\mathrm{h}, \mathrm{r}, \mathrm{t}$）$\in \mathrm{T}$，由此产生一个 KG 中不存在的三元组$G' = \{$（$\mathrm{h}, \mathrm{r}, \mathrm{t}$）$|\mathrm{h} \in$

$E, r \in R, t \in T,$（h, r, t） $\notin G\}$。知识图谱推理的目标是基于已有的知识，使用自动化方法推理得到潜在的实体之间的关系对和识别错误的知识，来完善知识图谱。比如知识图谱中已存在的信息是（阿里巴巴，总部，杭州）、（杭州，省会，浙江）、（浙江，省份，中国），通过推理可以得到（阿里巴巴，总部，中国）。知识推理的对象不仅包括实体之间的关系和属性名称，还包括实体的属性值和本体的概念层次。比如，如果已知一个人物实体的身份证号码，则由此可以知晓他的户籍地、年龄、性别等[130]。

无论知识图谱的规模多么庞大（可能存在数百万的实体和数亿的关系事实），总会有一些需要的事实在其中是缺失的。因此从已有的知识图谱中推理出新事实的能力很早就得到了研究者的重视，被称为知识图谱推理。知识图谱推理主要指利用知识图谱中已经存在的实体、关系或事实信息，来预测实体之间新关系的计算过程，在知识图谱的缺失信息补全、质量优化，以及增强问答系统和精准推荐等各项任务中具有重要意义。

2.2.2 知识图谱推理关键技术

AMiner

知识图谱本质上是一种语义网络，可以形式化地解释现实世界中的概念及其关系。知识图谱在结构化知识表达方式上不采用框架、脚本等繁琐的结构，而是使用灵活的三元组形式。因此，知识图谱推理不仅局限于传统的基于逻辑和规则的推理方法，还可以包含其他方法。近年来，开放域信息抽取技术的发展，为构建知识图谱提供了丰富的数据资源，多样的实例内容为知识推理技术带来了新的机遇和挑战。随着知识表示学习、神经网络等技术的发展，出现了一系列新的推理方法。

早期的知识图谱推理方法都是基于符号化的描述逻辑和规则。描述逻辑的显著优点是具有推理机制，能实现知识之间的自动推理。推理规则是可解释的，可以提供对推理结果的洞察。符号化的推理规则还能和机器学习结合来处理不确定性，被称为统计关系学习。许多利用神经网络来学习一阶逻辑规则的方法也已经被提出。虽然逻辑规则易于理解，但是它们对噪声敏感因此泛化性能差，后来被基于分布式向量表示的方法所代替。基于分布式向量表示的知识图谱推理方法又

被称为知识图谱嵌入。在知识图谱嵌入中，实体和关系被表示成隐空间中的连续向量。基于连续向量，多种评分函数被定义来计算一个三元组（e_s, r, e_o）。虽然知识图谱嵌入方法已经在几个知识图谱嵌入的数据集上取得了优秀的结果，一些研究显示它们在建模多跳关系时容易产生巨大的错误。而多跳关系对于更复杂的推理任务来说是不可避免的。除此之外，因为这些方法都在隐空间上操作，所以它们的预测是不可解释的。最近的工作将多跳推理和分布式表示结合起来，用深度学习来显式建模多步路径。这种方法能够同时享有分布式表示的泛化能力和逻辑规则的可解释性[131]。

基于资料调研结果，本章节通过详细介绍基于描述逻辑推理的方法、基于规则推理的方法、基于分布式表示推理的方法、基于神经网络推理的方法等推理方法的基本思想和实现原理，一方面帮助读者了解知识图谱推理相关技术，另一方面为读者在知识图谱推理领域的研究提供启发和思路。

（1） 基于描述逻辑推理的方法

AMiner

描述逻辑（Description Logic）是一种面向对象的知识表示的形式化方法，是一阶谓词逻辑的可判定性子集。与其他知识表示方法相比，描述逻辑的显著优点是具有推理机制，能实现知识之间的自动推理，因此描述逻辑近年来成为人工智能领域的研究热点。本节通过调研相关文献对描述逻辑及其推理方法进行介绍[132]。

一个描述逻辑系统包含 4 个基本组成部分：最基本的元素（概念、关系、个体）、TBox 术语集（概念术语的公理集合）、ABox 断言集（个体的断言集合）、TBox 和 ABox 上的推理机制。其中，针对最基本的元素，概念是指一个领域的子集（比如$x|\text{student}(x)$）；关系是指领域上的二元关系（比如$< x,y > |\text{friend}(x,y)$）；个体是指一个领域内的实例（比如：Alice）。TBox 为概念术语集，是描述概念和关系的知识，被称之为公理（Axiom）。TBox 语言有定义和包含，其中定义为引入概念及关系的名称，如 Mother、Person、has_child；包含指声明包含关系的公理，例如$c_1 \equiv c_2$。ABox 包含外延知识（又称为断言（Assertion）），来描述特定个体。ABox 语言包含概念断言和关系断言，概念断言是表示一个对

象是否属于某个概念，例如 Mother（Alice）、Person（Bob）；关系断言表示两个对象是否满足特定的关系，例如 has_child（Alice, Bob）。描述逻辑的语言包含三部分：构造算子、语法、语义，如图 26 所示。描述逻辑根据构造算子，在简单的概念和关系上构造出复杂的概念和关系。

构造算子	语法	语义	例子
原子概念	A	$A^I \subseteq \Delta^I$	Human
原子关系	R	$R^I \subseteq \Delta^I \times \Delta^I$	has_child
对概念C，D和关系(role) R			
合取	$C \sqcap D$	$C^I \cap D^I$	Human ⊓ Male
析取	$C \sqcup D$	$C^I \cup D^I$	Doctor ⊔ Lawyer
非	$\neg C$	$\Delta^I \backslash C$	¬Male
存在量词	$\exists R.C$	$\{x \mid \exists y.\langle x,y\rangle \in R^I \wedge y \in C^I\}$	∃has_child.Male
全称量词	$\forall R.C$	$\{x \mid \forall y.\langle x,y\rangle \in R^I \Rightarrow y \in C^I\}$	∀has_child.Doctor

图 26 描述逻辑的语义表

AMiner

基于描述逻辑推理中与知识图谱有关的推理手段是本体推理。本体（Ontology）作为一种结构化的知识表示形式，提供了不同领域的共享词汇，给计算机处理互联网信息提供了方便。OWL（Web Ontology Language）是 W3C 制定的最规范、最严谨的网络本体语言标准。OWL 的逻辑基础是描述逻辑，可以显示表达本体中词汇的确切含义、揭示其间的语义关系，从而形成对领域知识的共同理解，并通过一定的限制与规则使本体具有推理能力。图 27 给出了描述逻辑与 OWL 词汇的对应表。

Axiom	DL Syntax	Example
subClassOf	$C_1 \sqsubseteq C_2$	Human ⊑ Animal ⊓ Biped
sameClassAs	$C_1 \equiv C_2$	Man ≡ Human ⊓ Male
subPropertyOf	$P_1 \sqsubseteq P_2$	hasDaughter ⊑ hasChild
samePropertyAs	$P_1 \equiv P_2$	cost ≡ price
sameIndividualAs	$\{x_1\} \equiv \{x_2\}$	{Prsident_Bush} ≡ {G_W_Bush}
disjointWith	$C_1 \sqsubseteq \neg C_2$	Male ⊑ ¬Female
differentIndividualFrom	$\{x_1\} \sqsubseteq \neg\{x_2\}$	{John} ⊑ ¬{Peter}
inverseOf	$P_1 \equiv P_2^-$	hasChild ≡ hasParent$^-$
transitiveProperty	$P^+ \sqsubseteq P$	ancestor$^+$ ⊑ ancestor
uniqueProperty	$\top \sqsubseteq \leq 1P$	$\top \sqsubseteq \leq$ 1hasMother
unambiguousProperty	$\top \sqsubseteq \leq 1P^-$	$\top \sqsubseteq \leq$ 1isMotherOf$^-$

图 27 描述逻辑与 OWL 词汇的对应表

本体推理任务中比较常见的方法包括：基于 Tableaux 运算的方法、基于逻辑编程改写的方法、基于一阶查询重写的方法、基于产生式规则的方法，本节分别介绍这 4 个方法的基本思想、实现原理和相关工具。

- **基于 Tableaux 运算的方法**

基于 Tableaux 运算的方法适用于检查某一本体的可满足性，以及实例检测。其基本思想是通过一系列规则构建 ABox（个体的断言集合），以检测本体的可满足性，或者检测某一实例是否存在于某个概念，这种思想类似于一阶逻辑的归结反驳。Tableaux 的运算规则如图 28 所示，初始情况下，将 ABox 设置为$\emptyset$。第一条规则是指如果C、D（x）的交集是$\emptyset$，同时C（x）、D（x）都不在$\emptyset$中，那么可以理解为$\emptyset$只包含了 C 的一部分（C（x）不在里面），因此就把C（x）、D（x）都添加到$\emptyset$中。

$\sqcap^{+}$-规则：若$C \sqcap D(x) \in \emptyset$，且$C(x), D(x) \notin \emptyset$，则$\emptyset := \emptyset \cup \{C(x), D(x)\}$;
$\sqcap^{-}$-规则：若$C(x), D(x) \in \emptyset$，且$C \sqcap D(x) \notin \emptyset$，则$\emptyset := \emptyset \cup \{C \sqcap D(x)\}$;
$\exists$-规则：若$\exists R.C(x) \in \emptyset$，且$R(x,y), C(y) \notin \emptyset$，则$\emptyset := \emptyset \cup \{R(x,y), C(y)\}$,
其中，y是新加进来的个体;
$\forall$-规则：若$\forall R.C(x), R(x,y) \in \emptyset$，且$C(y) \notin \emptyset$，则$\emptyset := \emptyset \cup \{C(y)\}$;
$\sqsubseteq$-规则：若$C(x) \in \emptyset, C \sqsubseteq D$，且$D(x) \notin \emptyset$，则$\emptyset := \emptyset \cup \{D(x)\}$;
$\bot$-规则：若$\bot(x) \in \emptyset$，则拒绝$\emptyset$;

图 28 Tableaux 的运算规则

Tableaux 的相关工具包括 FaCT++、Racer、Pellet 和 HermiT，这些工具的相关介绍如图 29 所示。

工具名称	支持本体语言	编程语言	算法
FaCT++	OWL DL	C++	tableau-based
Racer	OWL DL	Common Lisp	tableau-based
Pellet	OWL DL	Java	tableau-based
HermiT	OWL 2 Profiles	Java	hypertableau

图 29 Tableaux 的相关工具简介

● **基于逻辑编程改写的方法**

由于基于 Tableaux 运算的本体推理方法，仅支持预定义的本体公理上的推理，无法针对自定义的词汇支持灵活推理，从而导致用户无法定义自己的推理过程。因此在基于逻辑编程改写的方法中引入规则推理，它可以根据特定的场景定制规则，以实现用户自定义的推理过程。该方法是面向 Datalog 语言，它是面向知识库和数据库设计的逻辑语言，表达能力与 OWL 相当，具有支持递归、便于撰写规则、实现推理等优点，可以结合本体推理和规则推理。

Datalog 是规则的集合，基本语法包含：原子（Atom）、规则（Rule）、事实（Fact）。原子的表示形式是p（$t_1, t_2, \ldots, t_n$），其中 p 是谓词，n 是数目，t_i是项（变量或常量），比如 has_child（x, y）；规则是由原子构建，它的表示形式是 $H\!:-B_1, B_2, \ldots, B_m$，其中 H 是头部原子，$B_i$是体部原子，比如 has_child（x, y）：－has_son（x, y）；事实是没有体部也没有变量的规则，表示形式是 F（$c_1, c_2, \ldots, c_n$），比如 has_child（x, y）：－。比如，规则集合是 $path(x, y)\!:-edge(x, y)$、$path(x, y)\!:-path(x, z), path(z, y)$，事实集合是 $edge(a, b)$、$edge(b, c)$，获得的结果是$path(a, b)$、$path(b, c)$、$path(a, c)$。

AMiner

基于逻辑编程改写的本体推理方法相关的工具包括 KAON2 和 RDFox，图 30 展示了它们支持的本体语言、编程语言等。

工具名称	支持本体语言	实现编程语言	支持编程语言
KAON2	OWL DL/SWRL	Java	Java
RDFox	OWL 2 RL	C++	Java/C++/Python

图 30 逻辑编程改写的相关工具简介

● **基于一阶查询重写的方法**

一阶查询是具有一阶逻辑形式的语言，因为 Datalog 是数据库的一种查询语言，同时具有一阶逻辑形式，因此可以使用 Datalog 作为中间语言，图 31 展示了处理流程，具体是：首先重写针对本体的查询语言 SPARQL 语言为 Datalog，然

后将 Datalog 重写为 SQL 查询，从数据库中获取数据，经过转换得到本体结果。基于一阶查询重写的方法，可以高效地结合不同数据格式的数据源，同时重写方法关联起了不同的查询语言。

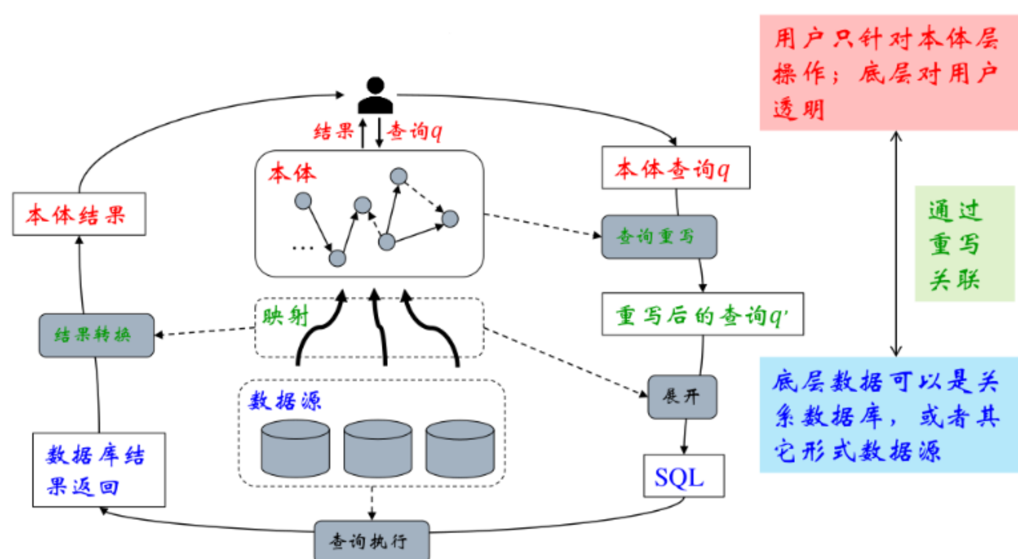


图 31 基于一阶查询重写方法的处理流程

AMiner

- **基于产生式规则的方法**

产生式规则系统是一种前向推理系统，可以按照一定机制执行规则从而达到某些目标，被应用于自动规划、专家系统上。产生式规则系统由：事实集合（Working Memory, WM）、产生式/规则集合（Production Memory, PM）、推理引擎组成，执行流程如图 32 所示。

事实集合用于存储当前系统中所有事实，包含描述对象和描述关系。描述对象的表示形式是（type $\mathrm{attr}_1{:}\,val_1\ \ attr_2{:}\,val_2 \ldots{:}\ \ attr_n{:}\,val_n$），其中 type、$\mathrm{attr}_i$、$\mathrm{val}_i$ 均是原子，比如（student name: Alice age:24）；描述关系是指（older Than John Alice）。产生式集合中的语句形式是“IF conditions THEN actions”，其中 conditions 是由条件组成的集合（简称 LHS），各条件之间是“且”的关系，当 LHS 中所有条件均被满足，则该规则被触发；actions 是由动作组成的序列（简称 RHS），各个动作是依次执行的，动作种类包括 ADD、REMOVE、MODIFY。比如 IF （Student name:x） Then ADD （Person name:x）。事实集合和产生式/规则集合是用户自定

义的数据，相当于 ABox 和 TBox。推理引擎可以控制系统的执行，包含模式匹配（用规则的条件部分匹配事实集中的事实，整个 LHS 都被满足的规则被触发，并被加入议程（agenda））、解决冲突（按一定的策略从被触发的多条规则中选择一条）、执行动作（执行被选择出来的规则的动作序列，从而对事实集合进行一定的操作）。

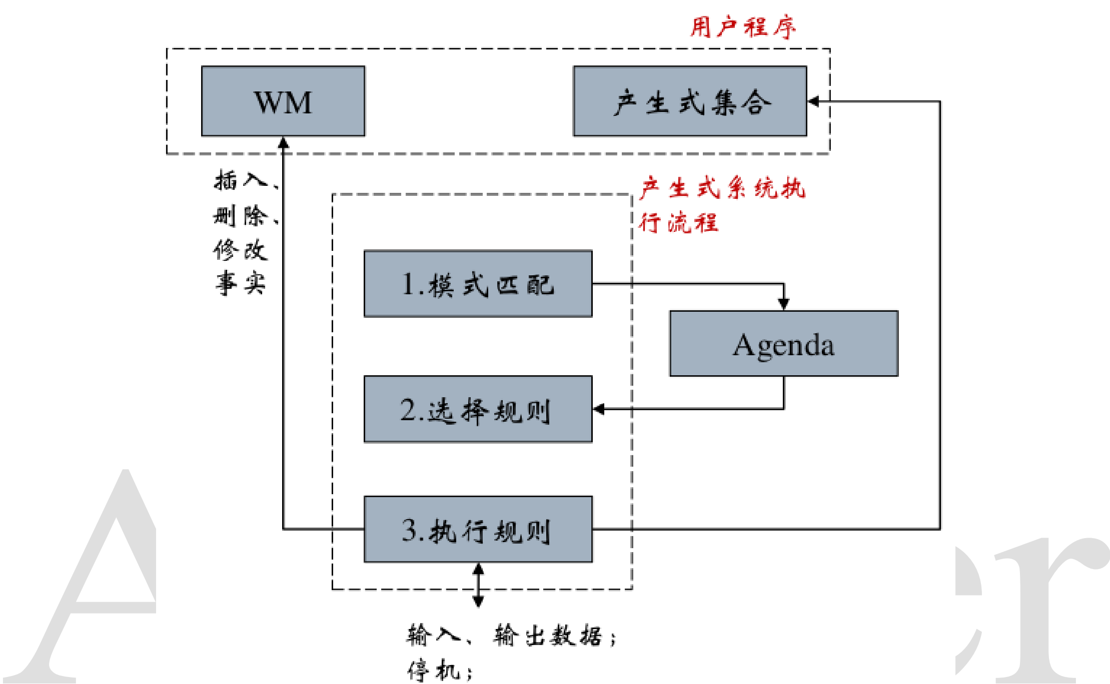


图 32 产生式规则系统的执行流程

基于产生式规则的方法主要应用到 Drools、Jena、RDF4J、GraphDB 等系统框架中，它们的介绍信息如图 33 所示。

工具名称	实现语言	支持编程语言	是否开源	算法
Drools	Java	Java	✓	RETEOO/PHREAK
Jena	Java	Java	✓	RETE
RDF4J	Java	Java/PHP/Python	✓	RETE
GraphDB	Java	Java/PHP/Python/Scala…	✗	TREE

图 33 产生式规则方法的相关工具

（2） 基于规则推理的方法

逻辑规则是知识图谱推理任务的有用表示，因为它们是可解释的，可以提供

对推理结果的洞察。在许多情况下，这种可解释性导致了传输任务的健壮性。图 34 展示了知识图谱推理任务中的逻辑规则示例[133]。规则“Has Office In Country”对于将新的公司和地点加入到知识图谱场景中也适用，但是知识图谱中实体嵌入表示方法（比如 TransE）处理该场景问题难以达到理想效果。

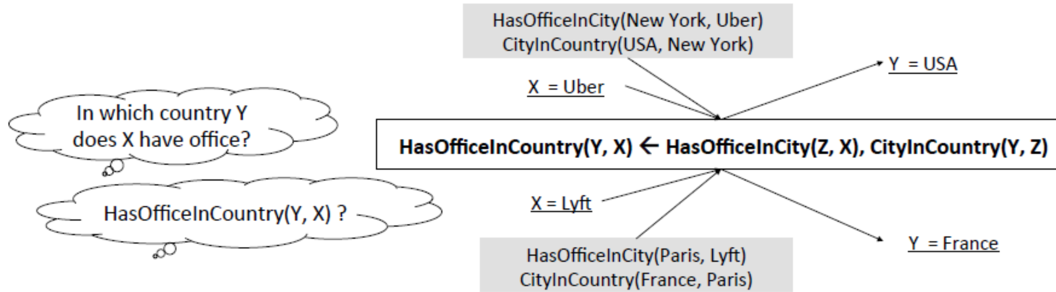


图 34 使用逻辑规则用于知识图谱推理任务的示例

关系规则集合学习是一种统计关系学习，当学习涉及到提出新的逻辑规则时，通常称为归纳逻辑规划。底层逻辑通常是概率逻辑，如马尔科夫逻辑网络或一阶概率语言模型 ProPPR。使用概率逻辑的优点是，通过为逻辑规则分配概率，可以更好地建模统计复杂和有噪声的数据。不幸的是，这个学习问题相当困难——它需要学习结构（即模型中包含的特定规则集）和参数（即与每个规则相关联的置信度）。确定结构是一个离散的优化问题，它涉及到对一个可能很大的问题空间的搜索。因此，过去的许多学习系统都使用了将离散结构空间中的移动与参数空间中的移动交错的优化方法。

考虑知识图谱结构的规则推理方法是在知识图谱上进行路径挖掘，将一些路径近似看成规则，通过实体间的路径来判断实体间是否存在指定关系的特征训练模型。这类工作代表性的方法是基于随机游走的规则挖掘、关联规则挖掘。基于随机游走的规则挖掘方法将每种不同的关系路径作为一维特征，通过在知识图谱中统计大量的关系路径构建关系分类的特征向量，建立关系分类器进行关系抽取。由于知识图谱的知识不够完善，这类方法存在数据稀疏问题。基于关联规则挖掘方法的目标是生成边关系规则，事先依据边类型生成所有可能的规则，再在图谱中找出支持该规则的事实，置信度达到阈值则认为该规则成立。这类方法也受限于数据稀疏的限制，很难实用。

AMiner

通过调研相关资料[134]，总结基于规则的知识图谱推理方法。NELL 知识图谱使用一阶关系学习算法进行推理，首先学习概率规则，然后人工筛选规则并根据具体的实体来实例化规则，最后基于规则从已有的关系实例推理得到新的关系实例。文献[135]使用一阶概率语言模型（programming with personalized PageRank，ProPPR）进行知识图谱推理。ProPPR 模型构建有向证明图，节点对应“关系（头实体变量，尾实体变量）”形式的子句的连接或推理目标。其中，起始节点为查询子句，边对应规则，也即一个推理步，从一个子句归约到另一个子句。边的权重与特征向量相关联，当引入一个特征模板时，边的权重可以依赖于模板的部分实例化结果，如依赖于子句中某个变量的具体取值。同时，在图中添加从每个目标尾节点指向自己的自环以及每个节点到起始节点的自启动边。自环用于增大目标尾节点的权重，自启动边使得遍历偏向于增大推理步数少的推理的权重。最后，ProPPR 中的推理基于图上的个性化网页评分实现。文献[136]进一步提出了 TensorLog，用可微的过程进行推理。TensorLog 中每个实体关联一个 one-hot 向量，为每个关系定义一个{0,1}操作矩阵，如果第 i 个实体和第 j 个实体存在对应的关系，（i, j）位置上的值为 1，否则为 0。这些表达都固定，不更新。由此逻辑规则推理可以形式化为矩阵相乘。给定实体和关系，预测另一个实体时，对于每一条可能的路径，实体 one-hot 向量乘以路径上关系操作的乘积（给定尾实体时是乘积的转置），用待学习的置信度参数对所有路径的结果进行加权求和，即可得到所有实体对应的得分向量。由于采用 one-hot 表示，候选实体的得分可以通过该得分向量乘以其 one-hot 向量的转置得到。对于置信度参数，通过最大化知识图谱中三元组的得分学习。文献[133]针对知识图谱推理任务，受到 TensorLog 的可微逻辑操作序列的启发，提出了一种将一阶逻辑规则的参数学习和结构学习相结合的端到端可微模型的神经逻辑规划框架，将推理任务编译成可微的数值矩阵序列，学习由一阶逻辑规则集合组成的模型，允许使用基于梯度的优化方法和目标规划。

文献[137]提出了 PRA 算法（Path Ranking Algorithm），将路径作为特征，预测实体间是否存在指定关系。PRA 首先确定要学习的目标关系；然后找出目标关系的正例三元组，替换头/尾实体得到负例三元组；再构造特征集合，将这些三元组中两个实体之间的一条路径作为一个特征；接着根据随机行走的思想计算路径的

AMiner

特征值，构成每个三元组的特征向量，每维对应一个特征的特征值；最后用这些正负例三元组对应的特征向量训练 logistic 回归分类器。文献[138]提出了知识图谱关联规则挖掘算法 AMIE（Association rule Mining under Incomplete Evidence），挖掘传递性规则。AMIE 维护一个规则队列，初始化为空规则，迭代地每次从规则队列中取出一个规则，如果规则是封闭规则（规则中的每个变量至少出现两次的规则）且未被删除，则输出该规则。同时，该规则用定义的规则挖掘操作算子（增加规则中的变量、实例化变量等）作用，将未被删除的规则插入规则队列。如此循环，直至规则队列为空。这里的删除条件是：如果$B_1 \wedge \ldots \wedge B_n \wedge B_{n+1} \Rightarrow H$的置信度小于$B_1 \wedge \ldots \wedge B_n \Rightarrow H$，即长规则并且置信度不够高将被丢弃。由于知识图谱的开放世界假设（open world assumption, OWA），不在知识图谱中的元组不能看成是负例，算法主要的挑战是提供负例。为此，算法提出了部分完整假设（partial completeness assumption，简称 PCA），如果知识图谱存在实体的某个属性，那么知识图谱存在该实体的所有属性。由此，置信度用 PCA 计算，算法可以进一步返回置信度超过阈值的规则。

虽然知识图谱上的推理规则通过算法自动挖掘，摆脱了人工挖掘规则成本高、覆盖率低的约束，但是自动挖掘的规则容易存在信息量低的噪音规则，从而误导推理，并且难以挖掘多样复杂的规则。基于规则推理的方法未来的研究方向是：引入局部结构信息，探索复杂算法，以支撑自适应挖掘复杂规则和有效特征。

（3） 基于分布式表示推理的方法

基于分布式表示推理的方法是指将知识图谱上的事实三元组中实体和关系等元素表示为低维向量形式，基于向量表示形式执行推理预测操作。通过调研相关资料[139][130]，本节从基于距离模型的表示推理方法、基于张量/矩阵分解的表示推理方法、基于语义匹配模型的表示推理方法等 3 个方面来介绍基于分布式表示推理技术的研究现状。

- **基于距离模型的表示推理**

基于距离模型的表示推理方法是根据转移假设（将关系看成实体间的转移）设计得分函数，度量事实三元组的有效性分值，分值越高表示该元组的有效性越

AMiner

高（即正例元组的分值高，而负例元组的分值低）。由于知识图谱中关系数量相对较少，负例常常通过替换头实体或尾实体得到（部分工作也替换关系构建负例）。推理预测时，选取与给定元素形成的元组得分高的实体/关系作为预测结果。TransE 是第一个基于距离的表示模型，主要思想是如果三元组（头实体，关系，尾实体）成立，头实体向量 h 与关系向量 r 的和与尾实体向量 t 相近，否则远离。得分函数是$-\|r+h-t\|_{L_1/L_2}$，用 L1 或 L2 范数度量距离，通过最小化一个基于 Margin 的损失，使得正例元组的分值比负例至少高一个 Margin。

TransE 模型可以很好的处理一对一关系，但是在处理一对多、多对一和多对多等复杂关系时存在一些不足和缺陷：（1）在处理多映射属性关系时，存在多个实体竞争一个点的现象；（2）未考虑知识图谱中的层级关系；（3）未考虑知识的丰富语义信息；（4）只支持在单个知识图谱上学习推理；（5）未考虑知识的时间约束。

针对以上问题，研究者们提出了一系列改进工作。文献[140]提出了 TransR 模型，在单独的实体空间和关系空间建立实体和关系的表示，每个关系对应一个空间，有一个映射矩阵，实体通过对应的映射矩阵映射到关系空间后，将关系向量看成实体向量间的转移，在一定程度上缓解了不能很好地处理多映射属性关系的问题。文献[141]提出的 TransG 模型是一种基于贝叶斯非参的高斯混合模型，第一次从产生式的角度看待知识图谱嵌入表示问题，解决了如何刻画三元组知识中关系的多语义表达问题。为了刻画关系的多语义性问题，TransG 模型的出发点是一个关系应该有多种向量表示，不同的实体对在几何变换中应该采用不同的关系向量，因此认为关系向量由若干子成分向量合成，模型会根据实体对自动选择一个关系向量，而多少个这样的关系向量还可以由模型自动选择。文献[142]提出了 TKRL（type-embodied knowledge representation learning）学习知识图谱实体和关系的表示，将层级类型信息用于映射矩阵、训练时负例的选择和评估时候选的过滤。文献[143]提出了 TransE-NMM（TransE-neighborhood mixture modeling），在 TransE 的基础上定义基于邻居的实体表示，引入邻居实体信息进行实体和关系的表示学习。文献[144]提出了 TEKE（Text-Enhanced Knowledge Embedding），引入文本语料中丰富的上下文信息扩展知识图谱的语义结构，学习知识图谱实体和

AMiner

关系的表示。文献[145]提出新颖的时间感知知识图谱补全模型 TAE（Time-Aware Embedding），用三元组和时间信息预测知识图谱中的连接，即：给定三元组中的两个元素与时间区间，预测另一个元素。

- **基于张量/矩阵分解的表示推理**

基于张量/矩阵分解的表示推理方法是将知识图谱事实三元组（头实体，关系，尾实体）作为元素构建张量/矩阵，通过张量/矩阵分解方法来推断未知的知识。张量/矩阵分解是将高维数组分解为多个低维矩阵的过程。分解得到的向量表示相乘重构成张量/矩阵，元素值即为对应三元组有效与否的得分，可以认为得分大于特定阈值的三元组有效，或候选预测按照得分排序，选择得分高的候选作为推理结果。一个三阶张量 X 中，两个节点表示实体对，第三阶模态保持了它们之间的关系。$X_{ijk} = 1$表示事实三元组（第 i 个实体，第 k 个谓词，第 j 个实体）存在，$X_{ijk} = 0$表示该事实三元组不存在。

RESCAL 是该类方法的典型代表，模型框架如图 35 所示[34]。RESCAL 基于三阶张量进行表示学习，利用张量因子分解，考虑了二元关系型数据的固有结构，把二元关系建模成一个张量：一个张量的两种模式表示实体，一种模式表示关系，当两个实体之间存在关系时，张量的元素为 1，否则为 0。RESCAL 模型可以解决规范的关系学习任务，表达式是$X_k \approx \mathrm{A}R_kA^T$，$X_k$表示二元关系的张量表示形式，A 是一个 n 行 r 列的实体表示矩阵，R_k是第 k 种关系对应的一个不对称的 r 行 r 列的矩阵，A^T是实体表示矩阵的转置。如果三元组成立，则三阶张量上相应的元素值为 1，否则为 0。

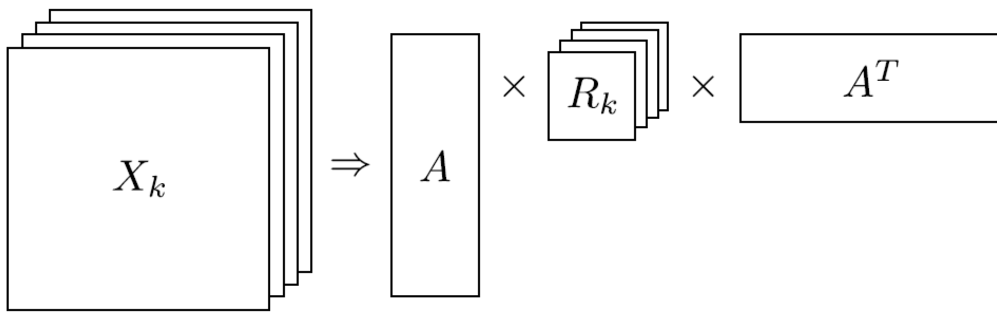


图 35 RESCAL 模型

RESCAL 模型虽然具有较高的推理准确率，但是内存占用量大，计算速度慢。为了解决以上问题，文献[146]提出了基于张量分解模型 RESCAL 的可扩展集成框架 RSTE（Random Semantic Tensor Ensemble）。RSTE 采用分治策略，从知识图谱中采样多个多样的更小规模子图张量，通过集成子图张量的 RESCAL 分解进行连接预测。RSTE 极大地降低了内存占用和运行时间，同时通过增加子图张量的分解或使某些子图张量分解得到的结果无效，可以快速处理动态变化知识图谱的增加或删除操作。文献[147]提出了新的张量分解模型 ARE（additive relational effects）学习知识图谱三元组的隐性和观察到的模式，用一个附加项增广 RESCAL 模型（隐性模式），对应观察到的模式。这里，观察到的模式是指用可观察的关系学习方法，例如规则方法等，得到的预测结果构成的三阶张量。附加项为该三阶张量乘以一个权重向量，权重向量衡量关系学习方法对各个关系的预测能力。该附加项通过减少不连接部分，降低 RESCAL 分解需要的阶。

- **基于语义匹配的表示推理**

AMiner

结构化的嵌入表示方法使用两个独立的矩阵来为每一个关系投射头实体和尾实体，这不能有效地表示实体和关系之间的本质联系。语义匹配模型首先将实体和关系分别表示为向量形式，然后对实体和关系之间的关系建模为语义匹配能量函数，并定义了语义匹配能量函数的线性形式和双线性形式。

文献[148]提出了全息嵌入（holographic embeddings, HOLE）来学习知识图谱组成成分的向量空间表示。HOLE 使用循环关联来生成知识图谱成分表示。通过使用相关性作为组合操作符，HOLE 可以捕获丰富的交互作用，同时保持推理的效率并易于训练。当前基于表征的关系推理模型存在的主要问题是忽略了实体和关系的语义多样性，从而制约了推理能力。针对以上问题，文献[149]提出了知识图谱中关系推理的一种新假设，即每个关系反映了相应实体的某些特定方面的语义联系，可以通过有选择地加权嵌入的组成部分来建模，以帮助缓解语义解析问题。在此基础上，提出了一种语义方面感知的关系推理算法，有效地提高了知识图上关系推理的准确性。文献[150]从类比推理的角度研究知识推理的解决方案，他们制定了类似的结构，并在评分函数中利用它们来优化实体和关系的潜在表示。

（4） 基于神经网络推理的方法

神经网络作为一种重要的机器学习算法，基本上是模仿人脑进行感知和认知。它在自然语言处理领域得到了广泛的应用，并取得了显著的效果。神经网络具有很强的特征捕获能力，通过非线性变换将输入数据的真实分布从原始空间转换为另一个特征空间，并自动学习特征表示。因此，它适用于知识推理等复杂任务。

神经网络用于知识图谱推理任务由来已久。文献[131]模仿认知科学中的双过程理论，提出了一个基于认知计算的深度学习推理框架 CogKR，能够访问知识图谱来进行多跳的关系推理。具体来说，CogKR 由一个扩展模块和一个推理模块组成，通过协调两个模块来构建一张认知图谱，从而能够基于认知图谱上的子图而不是路径来进行推理，从而适应更加复杂的推理场景。通过这些模型间的动态交互和端到端训练，CogKR 能够将它们结合成一个统一的结构，并且联合优化它们来进行知识图谱推理。

在结构化嵌入表示模型中，两个实体向量的参数不相互作用。为了缓解距离模型的问题，文献[151]介绍了一种使用标准的非线性单层神经网络来隐式连接实体向量的单层模型，可用于推理两个实体之间的关系。然而，非线性只提供了实体向量之间的弱相互作用。为了发现不同实体之间的隐含关系，文献[151]介绍了一种神经张量网络模型（neural tensor network, NTN），使用双线性张量层来替代传统的标准线性神经网络层。NTN 模型通过将知识图谱中每个实体表示为向量，来获取该实体的相关事实以及它隶属于某种特定关系的概率，通过利用神经张量网络的参数来定义每个关系，将关系表示为三阶张量，并明确地关联两个实体向量，由此能够在使用知识图谱现有知识的条件下准确预测其他事实。随着知识图谱规模的增加，知识推理模型的特征空间也越加复杂，为了简化模型结构，文献[152]提出了一种参数规模小但效果不错的投影嵌入模型（embedding projection, ProjE），该模型将实体预测看做候选项排序问题，通过将每个候选对象投影到一个由输入向量通过预先定义的组合运算符生成的目标向量上来计算投影相似度，选择排序分数最高的候选对象作为实体预测结果。大部分知识表示模型（如 NTN 模型、ProjE 模型）仅利用实体和关系名称的嵌入表示，而忽略了实体的描述信息，这些信息中存在丰富的语义信息可以作为推理的依据。文献[153]利用实体描

AMiner

述信息，提出了一种表示学习模型（description-embodied knowledge representation learning, DKRL）进行知识图谱中实体和关系预测。该模型使用两种编码器，包括连续词袋和深度卷积神经网络，通过学习实体的描述内容，不仅可以获得三元组的结构信息，还可以获得实体描述内容中的关键词和隐藏在语序中的文本信息。

部分研究者在循环神经网络的基础上，也提出了适合知识图谱推理的方法。文献[154]提出了一种利用循环神经网络（RNN），以非原子方式推理多跳连接关系的方法（简称 Path-RNN）。Path-RNN 使用路径排序算法为每个关系类型寻找不同的路径，然后将二元关系的嵌入表示作为输入向量。它在路径的第一个实体和最后一个实体之间关系的语义邻域中输出一个向量。三元组不是自然语言，用一个固定的表达式（h, r, t）来模拟复杂的结构。这样的短序列可能代表性不足，无法提供足够的信息进行推理。同时，从大量的路径中构造有用的长序列是昂贵和困难的。为了解决以上问题，文献[155]提出了 DSKG 模型，使用多层 RNN 来处理实体和关系。具体来说，DSKG 将独立的 RNN 单元用于实体层和关系层，因为这种为知识图谱特别设计的结构能够在关系多样化和复杂的情况下获得更好的性能。此外，DSGE 模型不仅可以预测实体，还可以推断三元组。文献将符号逻辑推理的丰富多步推理与神经网络的泛化能力相结合，使用 RNN 模型解决文本和大规模知识库中实体和关系的复杂推理问题。本文提出了建模的三个重要进展：（1）联合推理关系、实体和实体类型；（2）利用神经网络注意力机制建模融合多种路径；（3）使用单一 RNN 模型代表所有关系之间的逻辑组成，并在其中分享力量[156]。

在使用知识图谱来辅助问答系统时，我们有时只使用知识图谱中一个三元组来回答问题。然而，当问题复杂且知识图谱不完整时，问答系统必须能够利用现有的三元组来推断未知的答案。推理答案的过程可以建模为一个串行化的决策问题，因而可以通过强化学习来解决。文献[156]提出了一种新颖的强化学习框架（DeepPath）来执行知识图谱推理任务，该方法是首次将强化学习方法用于解决多跳推理问题。在 DeepPath 框架中，强化学习环境被建模为马尔可夫决策过程，使用元组<S, A, P, R>来表示。DeepPath 使用 TransE 和 TransH 来获得实体和关系的表示向量，知识图谱中所有关系定义了行动空间，两个实体的不同嵌入表示定

AMiner

义了状态向量。文献[158]提出了一种神经强化学习方法MINERVA，它学习如何根据输入查询引导图形找到预测路径，并避免了建模要求。MINERVA 将环境表示为一个确定性的、部分观察的马尔可夫决策过程，使用长短期记忆网络对其过去所做决策的历史进行编码，因而可以学习长链的推理。

基于神经网络的推理方法利用神经网络强大的学习能力来表示知识图谱中的三元组，从而获得更好的推理能力。然而，在知识图谱推理任务中，神经网络模型难以解释的问题仍然存在，如何解释神经网络的推理能力值得研究。迄今为止，基于神经网络的推理方法研究较少，但其强大的表现能力和在其他领域的突出表现，使其具有广阔的发展前景。在未来研究工作中，如何将现有的神经网络方法扩展到知识图谱推理领域值得探索。

2.2.3 知识图谱推理应用

知识图谱推理方法从已有的事实三元组中推断出未知关系，不仅为大规模异构知识图谱中的资源提供了高效的关联发现能力，而且在知识图谱中的知识验证和补全方面具有重要作用。

为了构建大规模的知识图谱，会从开放多样的数据资源中抽取知识，由于数据来源的质量难以保证，以及抽取技术的不完善，容易抽取错误的知识，使得知识图谱内部存在噪声知识和知识矛盾现象。比如 NELL 知识图谱使用模板匹配技术从互联网中抽取知识，随着时间推移，采集到的知识的准确率不断下降。主要原因是抽取模板难以从存在噪音的数据中抽取准确的知识，而错误的知识会产生不可靠的模板，使用错误的模板会抽取的错误知识。如此循环迭代，NELL 的知识质量会大大降低，虽然会定期使用人工删除错误的事实元组，但是时间成本和人工成本较高，因此需要知识推理方法，自动对知识图谱中的知识进行验证，降低成本提高效率。此外，现有的知识图谱由于数据来源的不全面以及知识抽取技术的遗漏，难以构建信息完备的知识图谱，实体的属性和实体之间的关联关系等信息匮乏严重。目前大部分研究者使用知识推理方法，利用知识图谱中现有的知识推理得到隐含的知识来获得新知识，从而补全知识图谱。

推理技术可以通过对不同领域的知识和规则建模来进行领域知识推理，从而

AMiner

支持自动决策、数据挖掘和链接预测等功能，可以广泛应用到医疗、金融、智能问答系统、推荐系统等场景中。

目前，医学领域已经成为知识图谱被积极应用的领域，也是人工智能领域的研究热点。目前国内外医学知识图谱的研究都取得了很大进展，国内主要有康夫子“医疗大脑”和百度“医疗大脑”，二者重点关注对 3000 余种常见病的建模；国外比较著名的是 IBM Watson，用来辅助医生在肿瘤领域进行诊疗决策。知识图谱是智能医疗的底层核心，而基于知识图谱的知识推理是疾病诊断过程的基础。通过借助医学知识图谱，知识推理能够帮助医生根据病患的身体症状及近期摄入或接触物品等完成病患数据搜集与疾病诊断，从而为医生制定决策参考，降低医疗差错率。

金融也是使用知识图谱的一个活跃领域。知识图谱中的投资关系和雇佣关系可以通过聚类算法来识别利益相关者群体。当某些节点发生了更改或发生了大型事件时，可以通过路径排序和子图发现方法推断更改实体之间的关联。在金融行业，反欺诈是一项重要的工作。通过知识推理，人们可以对信息的一致性进行验证，提前识别欺诈行为。知识图谱推理提高了金融行业的资源配置效率，增强了风险管理和控制的能力，有效地促进了金融行业的发展。

知识图谱是智能问答系统的强大数据支撑，由于目前知识图谱的知识还不完备，需要推理技术的支持，代表性的系统包括 IBM 的 Watson、苹果公司的 Siri、亚马逊的 Alexa、百度的度秘等，它们基于知识图谱及推理技术，提供精确、简洁的答案。其中，美国知识竞赛节目“危险边缘”的问题涉及多个领域，要求参赛选手需要具备各个领域的知识，分析和推理其中包含的讽刺和谜语。IBM 的 Watson 在 2011 年的该节目中战胜了人类冠军选手 Ken 和 Brad。

基于知识图谱的推荐系统连接用户和项目，可以整合多个数据源，丰富语义信息。隐式信息可以通过推理技术获得，从而提高推荐的准确性。基于知识图谱推理方法进行推荐的典型案例有：购物推荐、电影推荐、音乐推荐等。知识推理能够更好地理解用户的搜索意图，提供接近“专、精、深”的垂直搜索，推荐符合用户心意的产品。

AMiner

2.2.4 知识图谱推理研究问题与挑战

前述内容中，我们总结概述了知识图谱推理的一系列技术方法，包括基于描述逻辑推理的方法、基于规则推理的方法、基于分布式表示推理的方法、基于神经网络推理的方法。这些推理方法既有区别又有相似之处，在推理任务中是互补的。它们都将知识图谱抽象成拓扑形式，然后利用实体之间的拓扑关系对特征进行建模和参数学习。主要区别在于：

✧ 基于描述逻辑和规则的推理方法可以捕获知识图谱中隐藏的语义信息，显著提高知识推理的准确性，并且可以模拟人类的推理能力，这使得利用先验知识来辅助推理成为可能。然而通过人工定义的逻辑规则存在成本高、覆盖率低的局限性，而通过算法自动挖掘的推理规则容易包含噪音，从而误导推理。此外，目前获取的规则形式比较单一，泛化能力差，随着知识图谱中的内容增加和丰富，对复杂规则的需求越来越大。

✧ 基于分布式表示推理的方法是将实体和关系映射到低维空间向量，利用语义表达式进行推理。优点是充分利用了知识图谱中的结构信息，方法简单容易扩展到大规模知识图谱中。缺点是这类方法在建立推理模型时没有考虑先验知识，只考虑满足知识图谱中事实三元组的约束条件，缺乏更深层次的成分信息，限制了推理能力。

✧ 基于神经网络的知识推理模型将 CNN 或 RNN 模型集成到表示学习模型或逻辑规则模型中，通过深度学习模型的自学习能力抽取特征，然后利用它的记忆推理能力建立实体关系预测模型。基于神经网络的推理方法直接对事实三元组建模，具有较强的学习能力、推理能力和泛化能力，但是模型复杂并且可解释性差。

下面总结知识图谱推理面临的几大问题与挑战：

（1） 知识图谱不完备。万事万物都处在一个复杂的因果网络中，知识图谱中信息大多是业务结果数据，缺乏产生这些数据的背景因果，比如常识知识。人具有推理能力主要得益于人人都知道、无需言明的常识知识，但是知识图谱构建技术在常识知识挖掘方面关注力不高，导致知识图谱中常识知识覆盖率低，难以

AMiner

支撑一些复杂推理。此外，知识图谱是不完整的，实体和关系等信息的缺失，导致利用实体描述信息或者关系路径的推理方法难以取得良好结果，功能拓展受到限制。

（2） 知识图谱质量难以保证。高质量的数据资源难以满足构建大规模知识图谱的需求，因此知识图谱构建技术开始关注互联网上公开的数据集，这类数据虽然规模很大，但是具有碎片性、多样性、不规范性、存在噪音等特点，再加上知识抽取技术的不完善，使得知识图谱中的知识质量，以及规范性方面都难以保证。而基于规则、分布式表示、神经网络的推理技术对噪音数据比较敏感，错误的或模棱两可的数据将导致学习偏差，降低了推理结果的准确性。

（3） 神经网络模型的黑盒问题。神经网络模仿人的大脑思维方法，采用自学习和自适应算法，能够从大规模数据中总结规律，完成运算、推理、识别等任务，表达能力丰富、推理能力强，并且容错能力高，但是神经网络模型具备黑盒特性，人们难以解释和表达神经网络模型得出结论的原理或原因，在与其他方法比较时难以获得更好理论支撑，从而导致不能广泛应用到其他任务问题。

AMiner

（4） 应用场景复杂。医学知识图谱的研究越来越成熟，支撑了大量的医学应用。但是由于医学领域的主观判断依赖特点（医生针对同一种疾病，会根据病人的情况作出不同诊断，），使得医学知识图谱中存在大量重复矛盾信息，增加了医学推理模型的复杂性。知识图谱推理提高了金融行业的资源配置效率，增强了风险管理和控制的能力，有效地促进了金融行业的发展。然而，由于金融行业数据的标准化程度较低，且分散在多个数据系统中，现有的数据分析和推理方法难以满足大规模数据分析的要求。

知识推理作为知识图谱的一个分支，在实现和应用上虽然具有很大技术难度，但是国内外的研究热度逐年上升，并不断取得新成果。知识推理在智能问答、深度搜索、辅助决策等领域展现出了重要作用，使智能体实现知识应用并与客观世界进行充分交互，是实现认知智能的关键环节。

2.2.5 知识图谱推理未来研究方向

人们在知识图谱推理方面进行了大量研究，现有模型也显示了它们的推理能

力，但是仍然有许多改进的地方有待探索。本节基于前述内容的分析结果，讨论知识图谱推理的未来研究工作。

（1） 动态知识推理。知识图谱不是静态的，它的内容会随着时间推移而发展，尤其有些信息具有时效性，比如国家领导人有任期约束，他们的头衔只在一定时间内有效。因此在推理过程中需要考虑时间信息，然而现有的大部分推理方法往往忽略了时间因素。因此，为了开展更深层次的研究，并提高知识推理结果的准确性，需要进一步探索动态知识图谱推理方法，灵活处理知识图谱上知识的增加、删除和修改等操作。

（2） 融合多源信息的知识推理。随着移动通信技术的快速发展，人们可以随时在网上上传和分享包括文本、音频、图像和视频在内的多媒体内容。如何高效、有效地利用这些丰富的信息成为一个关键和具有挑战性的问题。知识图谱内含信息丰富，既有语义信息，如实体名称、关系名称、实体描述、属性能信息，又包含大量的结构信息，如三元组、多元组、路径等，这些与事实三元组紧密相关的信息都可以作为推理的依据。多源信息已显示出在知识图谱上进行推理的潜力，然而利用这些信息的现有方法仍然处于初步研究阶段。我们可以设计更有效和优雅的模型来更好地利用这些信息。

（3） 融合多种方法的知识推理。知识图谱推理方法主要分为基于规则、基于分布式表示、基于神经网络等。不同的方法具有各自的优缺点，为了提高推理方法的性能，尝试将多种方法交叉融合，实现优势互补是未来值得研究的方向。比如，神经网络具有突出的学习能力和泛化能力等优势，但是因为黑盒操作，具有不可解释性等缺点，而规则方法具有较高的准确率和可解释性，如果将它们进行结合，可以获得极好的推理结果。

（4） 小样本的知识推理。现有的知识图谱推理模型往往需要大量的高质量样本进行训练和学习，同时也会耗费大量的时间和人力。在实际应用中，难以获得大量的高质量、覆盖面广的训练样本，极大地限制了现有知识推理模型（比如基于分布式表示的推理方法、神经网络推理模型）的应用范围。考虑到人脑具有极强的学习和推理能力，并且能够基于小样本数据学得大量有效知识。因此设计可以模拟并学习人脑获取知识、利用知识并实现推理等思维方式的推理方法，

AMiner

可以有效解决对大规模训练数据的依赖问题。

（5） 平衡推理方法的复杂性和灵活性。随着数据量的递增和业务需求的扩展，知识图谱的规模和内容复杂度越来越高，比如实体之间的多元关系、不同语言的知识版本。现有推理方法主要针对同一语言版本、实体之间的二元关系来设计的，容易导致忽略大量有用信息。相比其他语言，英文版的知识图谱内容覆盖率最高，在对其他语言的知识图谱进行推理时，可以将英文知识图谱作为推理依据。此外，知识图谱上大量实体具有多元关系（比如，Freebase 中超过三分之一的实体具有多元关系），如果将多元关系简化为二元关系处理容易丢失大量的语义信息。但是推理模型的表达能力和推理能力难以兼顾，表达能力越强，表达结构越复杂，灵活性越差，推理能力相应地越弱。如何权衡表达能力与推理能力，一方面尽量减少信息损失，另一方面使得推理能够灵活进行，是未来研究工作需要探讨的方向。

AMiner

2.3 逻辑表达

逻辑表达是一门研究思维形式结构及其基本规律的科学。逻辑基本规律是正确思维的根本假定，也是理性交流的必要条件。学习逻辑表达，有助于提高表达能力，使我们说话、写文章有条理，因为表达思想要靠语言，靠说话和写文章。话是说给别人听的，文章是写给别人看的，所以不仅需要自己理解，还需要对方理解。要把思想传达给别人，自己先要有明确的概念和恰当的判断，然后还要合乎逻辑地表达出来。如果表达不合逻辑，容易给别人理解表达内容造成困难。

自然语言生成是研究使计算机具有人一样的表达和写作的功能的技术，为了自动生成一段高质量的自然语言文本，必须具备良好的逻辑表达能力。因此本文着重介绍自然语言生成的概念、发展历程、关键技术、研究问题与挑战，以及未来研究方向。

2.3.1 自然语言生成概念

AMiner

自然语言生成（Natural Language Generation, NLG）是人工智能和计算语言学的分支，研究计算机能够根据一些关键信息及其在机器内部的表达形式，经过一个规划过程，来自动生成一段高质量的自然语言文本。随着自然语言处理研究热度的持续高涨，自然语言生成渐渐进入人们视野，并受到广泛关注。

自然语言生成和自然语言理解（Natural Language Understanding, NLU）都是自然语言处理（Natural Language Processing, NLP）的研究分支。NLU 的关注重心是以人类语言（自然语言）作为输入，经过处理后，输出机器可读的语义表示；NLG 主要关注将语义信息以人类可读的自然语言形式进行表达，基于语法和语义规则生成自然语言文本。二者的侧重点不同，NLU 实际上是使文本的结构和语义逐步清晰的过程，而 NLG 的研究重点是确定哪些内容是满足用户需要必须生成的，哪些内容是冗余的。虽然 NLU 和 NLG 的侧重点不同，但是二者存在诸多共同点：（1）研究基础都是语言模型；（2）需要使用语法规则，并解决省略、指代等语用问题。两者在实际使用过程中，是相辅相成的，存在于同一系统的不同阶段，发挥着不同作用。在做好自然语言理解的前提下，才能做好自然语言生成，进而产生很多落地的产品和应用。

自然语言生成的目标是生成高质量的人类可读的文本信息。根据生成文本的长短不同，NLG 分为句子生成和篇章生成。比如，在问答系统、对话系统等场景中，根据上下文生成句子；在写稿机器人、自动文摘、文案创作等场景中，生成的文本形式是段落。根据输入信息的不同，NLG 可以分为：数据到文本的生成、文本到文本的生成、图像到文本的生成。数据到文本的生成主要指根据输入的数据，获取对数据的文本分析内容，比如报告生成、新闻生成等；文本到文本是指输入的是文本，生成的也是文本，比如机器翻译、自动文摘等，这一方面的研究受到的关注度比较高；图像到文本的生成是指根据输入的图像，自动生成一段文字描述。

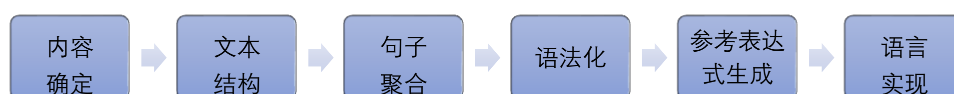


图 36 自然语言产生框架

为了跨越人类和机器之间的沟通鸿沟，将非语言格式的数据转换成人类可以理解的语言格式，自然语言生成系统包括 6 个步骤：内容确定、文本结构、句子聚合、语法化、参考表达式生成、语言实现，如图 36 所示。内容确定（Content Determination）作为第一步，需要 NLG 系统决定哪些信息应该包含在正在构建的文本中，哪些不应该包含，通常数据中包含的信息比最终传达的信息要多。确定需要传达哪些信息后，文本结构（Text Structuring）步骤需要合理的组织文本的顺序。例如在报道一场篮球比赛时，会优先表达“什么时间”“什么地点”“哪 2 支球队”，然后再表达“比赛的概况”，最后表达“比赛的结局”。不是每一条信息都需要一个独立的句子来表达，将多个信息合并到一个句子里表达可能会更加流畅，也更易于阅读。句子聚合（Sentence Aggregation）步骤是如何将多个信息合理并富有逻辑性的合并到一个句子中。当每一句的内容确定下来后，就可以将这些信息组织成自然语言了。语法化（Lexicalisation）步骤的工作是在各种信息之间加一些连接词，看起来更像是一个完整的句子。参考表达式生成（Referring Expression Generation）步骤跟语法化很相似，都是选择一些单词和短语来构成一个完整的句子。不过他跟语法化的本质区别在于“参考表达式生成需要识别出内容的领域，然后使用该领域（而不是其他领域）的词汇”。最后，当所有相关的单词和短语都已经确定时，语言实现（Linguistic Realization）步骤需要将它们组

合起来形成一个结构良好的完整句子[159][160]。

2.3.2 自然语言生成关键技术

自然语言生成是人工智能和计算机语言学的分支，同时作为理论语言学与计算机语言学共同的研究课题，在近年来渐渐成为人们关注的焦点，相应的自然语言生成技术也逐步从传统技术向新型技术迈进。通过调研相关资料，本节介绍了基于模板的自然语言生成技术和基于神经网络的自然语言生成技术的实现原理和研究现状。

（1） 基于模板的自然语言生成

模板生成技术是人们最早采用的一种自然语言生成技术，也是一种最简单的技术。这种生成技术的原理和填充方法相似，系统事先设计好几种可能出现的语言情况，构造几个相应的模板，每个模板包括一些常量和一些变量，它们往往由语言学家参与整理，在定义这些模板的结构时，应当力求让语言学家容易了解和书写这些模板。当用户输入一定的信息后，文本生成器将这些信息作为字符串嵌入到模板中替代变量。

AMiner

```
Topic->weather
  Act->query
    Content·:·weather_state
      ->3 对不起，请[<tell>]您需要[<refer>]{<where>}的[<what>]。
      ->2 请[<tell>]您需要[<refer>]的[<what>|具体内容]。
      ->1 抱歉，请[<tell>]您需要{<refer>}{(day)|今天|[when]} {(location)|当前城市
      |<where>}的[<what>]。

符号说明：
|：或者
[·]：内部元素出现次数≥1
{·}：内部元素出现次数≤1
(·)：对话管理模块的模板中的变量
<>：自定义语料中的变量
句子前的数字：该句子的权重，权重越大句子出现的可能性越大
```

图 37 询问天气场景中的句子模板

自然语言生成模板由句子（sentence）模板和词汇（word）模板组成。句子模板包括若干个含有变量的句子，词汇模板则是句子模板中的变量对应的所有可

能的值。为了方便理解，图 37 和图 38 分别提供了询问天气场景中的句子模板和词汇模板的可视化表示[161]。

<tell>->[告诉我|补充|说明|输入]
<refer>->[查询|知道|获取|收到|了解|咨询]
<where>->[哪里|何处|什么位置|什么地方|什么城市|哪个位置|哪个区域]
<what>->[天气|哪方面信息|什么信息|哪方面情况|哪方面内容|何种内容]
<when>->[哪天|什么时间|哪个时辰|什么时候]

图 38 询问天气场景中的词汇模板

一般的模板系统大多从实现的角度进行描述，它的处理只是在字符串的水平上，没有在深层次上进行语言处理，所以文本有一定的不完备性。这种技术虽然效率高、实现手段简单，但生成的文本质量不高，造成了模板描述语言对于用户（语言学家）而言不够自然的问题，影响了该类系统的可维护性和可扩充性。

由于自然语言生成模板思路简单，目前仍具有十分广泛的用途。比如，任务驱动的对话系统中的对话管理模块会根据当前的对话状态、用户输入等信息，产生下一步动作相关的信息，也就是会确定自然语言产生模块应该选择的句子模板和可选的词汇模板。为了提升自然语言生成模板生成的文本质量，满足人们多变的需要，用于自然语言生成的模板描述语言首先应当符合人的思维习惯，使得模板易于书写；其次用于自然语言生成的模板描述语言应当有较强的描述能力，能够表达尽可能多的语言现象；另外实际应用中需要模板描述语言易于扩充，以便在其中加入新的成分，使其能够描述现有语言成分无法描述的新的语言现象，保证模板描述语言的健壮性。

AMiner

（2） 基于神经网络的自然语言生成

随着深度学习的快速发展，各种深度学习模型在许多自然语言生成任务上取得不错的效果，尤其是近期的 GPT、BERT、XLNet 等预训练模型，大大提高了计算机处理自然语言任务的效果，比如机器翻译、文本摘要等。

本节对在自然语言生成领域广为应用的深度学习模型，包括 ELMO、Transformer、GPT、BERT、XLNet 等，详细介绍了这些模型的实现原理，以及各自的优缺点，以此为读者了解基于神经网络的自然语言生成方法提供信息窗口。

- **ELMO（Embeddings from Language Models）**

2018 年 3 月份，ELMo（Embeddings from Language Models）出世，提出该模型的论文 *Deep contextualized word representations*[162]是 NAACL18 Best Paper。ELMo 是一种新型深度语境化词表征，可对词进行复杂特征（如句法和语义）和词在语言语境中的变化进行建模（即对多义词进行建模）。词向量模型需要解决两个问题：（1）词使用的复杂特性，如句法和语法。（2）如何在具体的语境下使用词，比如多义词的问题。传统的词向量（如 word2vec、GloVe）模型中每个词对应一个向量，能够解决第一类问题，但是无法解决第二类问题。ELMo 的工作对于此，提出了一个较好的解决方案。使用时，将一句话或一段话输入模型，模型会根据上下文来推断每个词对应的词向量。这样做之后明显的好处之一就是对于多义词，可以结合前后语境对多义词进行理解。比如 apple，可以根据前后文语境理解为公司或水果。

AMiner

ELMo 主要使用了一个两层双向的 LSTM 语言模型，结合上下文来理解词义，模型结构如图 39 所示。ELMo 模型先在一个大型的预料库上面进行训练，用深度双向语言模型（biLSTM）内部状态的函数来表征一个词的向量，此时内部状态混合了所有的语义；然后将下游任务的文本输入模型，此时的词具备了具体的上下文语境，再用内部状态的线性组合来表征词。这种线性组合是在下游任务中进行训练得到的。

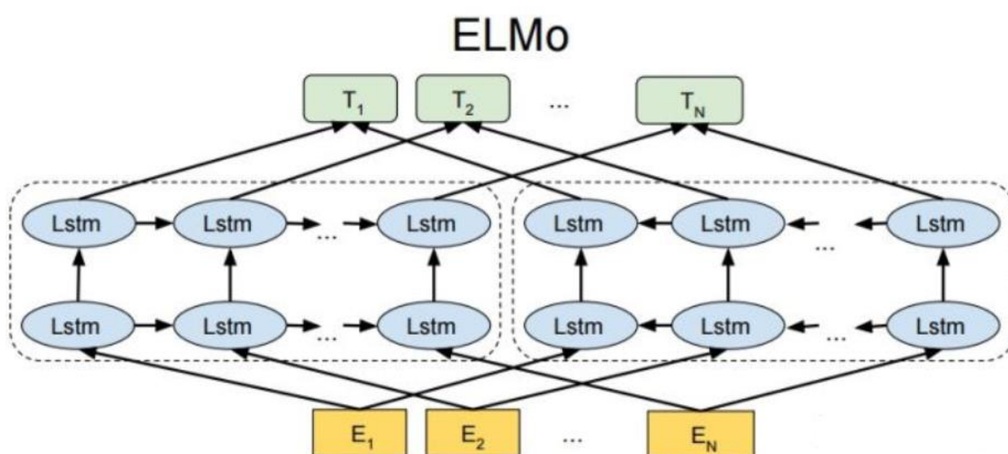


图 39 ELMo 模型架构

ELMo 模型中的 BiLSTM 层能有效地编码上下文中不同类型的语法和语义

信息，提高整体任务性能。但是 ELMo 模型采用 LSTM 作为特征提取器，其特征提取能力弱于 Transformer。

- **Transformer**

2017 年，Google 在 *Attention Is All You Need* 文章中提出了一种将 Attention 思想发挥到极致的模型 Transformer，摒弃了传统的 RNN/CNN 模型，主要基于 self-attention 机制设计，可以快速并行实现。Transformer 模型广泛应用于机器翻译，问答系统，文本摘要和语音识别等 NLP 领域。

Transformer 模型和 seq2seq 模型一样，也采用了 encoder-decoder 架构。但其结构相比于 Attention 更加复杂，encoder 层由 6 个 encoder 堆叠在一起，decoder 层也一样。Transformer 模型的结构框架如图 40 所示[163]。

Encoder 由 N=6 个相同的 layer 组成，layer 指的就是下图左侧的单元（即 Nx）。每个 layer 由两个 sub-layer 组成，分别是 multi-head self-attention mechanism 和 fully connected feed-forward network。其中每个 sub-layer 都加了 residual connection 和 normalisation，因此可以将 sub-layer 的输出表示为：

$$sub_layer_output = LayerNorm\left(x + \left(SubLayer(x)\right)\right)$$

AMiner

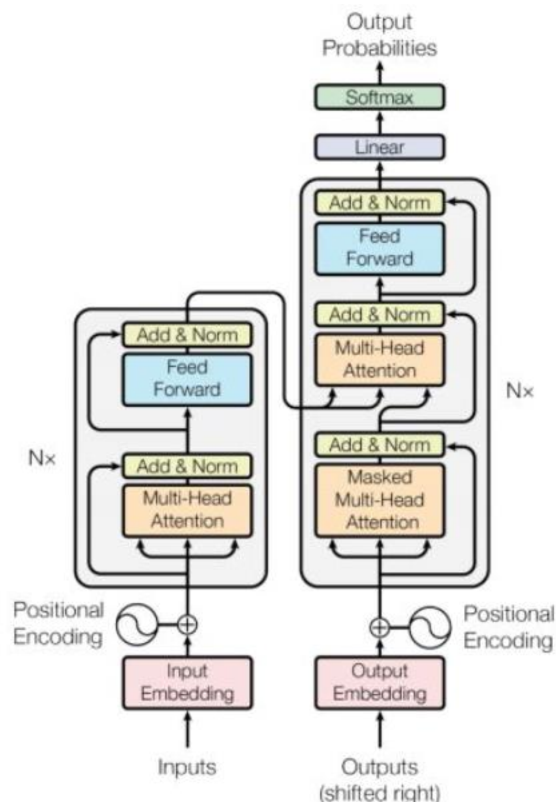


图 40 Transformer 模型架构

Decoder 和 Encoder 的结构差不多，但是多了一个 attention 的 sub-layer。Decoder 的主要不同是新加的 attention 多加了一个 mask，因为训练时的 output 都是 ground truth，这样可以确保预测第 i 个位置时不会接触到未来的信息。

Transformer 是第一个用纯 Attention 搭建的模型，不仅计算速度更快，在翻译任务上也获得了更好的结果。但是在实践上，有些 RNN 模型轻易可以解决的问题 Transformer 没做到，比如复制字符串，或者推理时碰到的序列长度比训练时更长（比如位置嵌入）。此外，在理论上，Transformer 模型不是图灵完备的，因而无法实现 while 循环。

- **GPT（Generative Pre-Training）**

2018 年，OpenAI 在论文 *Improving Language Understanding by Generative Pre-Training*[164]中提出了 GPT 模型。OpenAI GPT 是一类可迁移到多种 NLP 任务的，是一种基于 Transformer 的语言模型，利用 Transformer 的结构来进行单向语言模型的训练，基本思想是在尽量不改变模型结构的情况下将预训练的语言模型应用

到各种任务。OpenAI GPT 模型结构如图 41 所示。

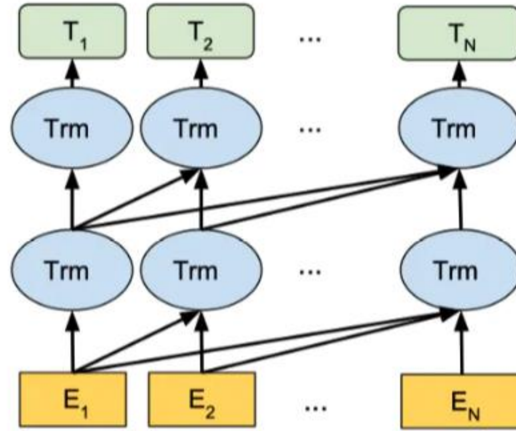


图 41 GPT 模型结构

GPT 预训练的方式和传统的语言模型一样，通过上文，预测下一个单词。GPT 因为采用了传统语言模型所以更加适合用于自然语言生成类的任务，因为这些任务通常是根据当前信息生成下一刻的信息。

GPT 训练过程分为两个部分，第一部分是利用没有标签的文本数据集训练语言模型，即：无监督预训练语言模型；第二部分是根据具体的下游任务，例如自动问答、文本分类等对模型进行微调，即：有监督的下游任务 fine-tuning。

AMiner

✧ 无监督预训练语言模型

给定句子语料$U = u_1, \dots, u_n$，目标函数为最大化似然函数：

$$L_1(U) = \sum_i logP(u_i|u_{i-k}, \dots, u_{i-1}; \theta)$$

该模型中应用 multi-headed self-attention，并在之后增加 position-wise 的前向传播层，最后输出一个分布。W_p 是单词位置的 Embedding，W_e 是单词的 Embedding。用 voc 表示词汇表大小，pos 表示最长的句子长度，dim 表示 Embedding 维度，则 W_p 是一个 pos×dim 的矩阵，W_e 是一个 voc×dim 的矩阵。得到输入 h_0 之后，需要将 h_0 依次传入 GPT 的所有 Transformer Decoder 里，最终得到 h_t。根据 h_t 再预测下个单词的概率 P（u）。

$$h_0 = UW_e + W_p$$

$$h_l = transformer_b lock(h_{l-1})$$

$$P(u) = softmax(h_n W_e^T)$$

✧ **有监督的下游任务 fine-tuning**

GPT 经过预训练之后，会针对具体的下游任务对模型进行微调。微调的过程采用的是有监督学习，训练样本包括单词序列$x^1, \ldots, x^m$和类标 y。GPT 微调的过程是根据$x^1, \ldots, x^m$和类标 y，通过语言模型得到h_l^m，经过输出层后对 y 进行预测：

$$P(y|x^1 , \ldots, x^m) = softmax\left(h_l^m W_y\right)$$

W_y表示预测输出时的参数，微调时需要最大化以下函数：

$$L_2(C) = \sum_{(x,y)} logP(y|x^1 \ldots, x^m)$$

GPT 在微调的时候也考虑预训练的损失函数，所以最终需要优化的函数为：

$$L_3(C) = L_2(C) + \lambda * L_1(C)$$

GPT 是基于 Transformer 的语言模型，相较于 RNN 能捕捉更长范围的信息，在计算速度上快于 RNN，易于并行化。但是 GPT 模型的可扩展性不强，对于某些类型的任务需要对输入数据的结构作调整。

- **BERT（Bidirectional Encoder Representations from Transformer）**

BERT 模型由 Google 在 2018 年发表的论文 *BERT: Pre-training of deep bidirectional transformers for language understanding*[165]中提出，在多项经典 NLP 任务中全面超越最佳模型，并为下游具体的 NLP 任务提供了极其简单的接口，是 NLP 领域里程碑式的成果。

AMiner

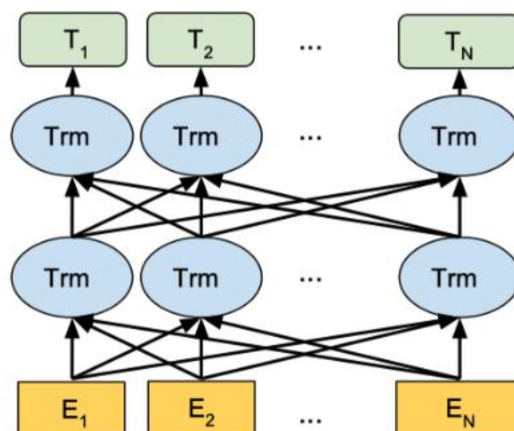


图 42 BERT 模型的架构

BERT 模型是 sentence-level 的词向量模型，充分描述字符级、词级、句子级甚至句间关系特征，增强了词向量模型泛化能力，结构如图 42 所示。BERT 的模型架构基于 Transformer，实现了多层双向的 Transformer 编码器。文中有两个模型，一个是 1.1 亿参数的 base 模型，一个是 3.4 亿参数的 large 模型。其中 base 模型的参数和 OpenAI 的 GPT 的参数一致，目的就是为了同 GPT 的效果进行一个比较。

AMiner

BERT 模型有三个创新点：bidirectional transformer、masked language model、sentence-level relation。

- bidirectional transformer：当理解语言时，前序词语会对后序词语的理解有帮助；同样地，后续词语也对前向词语的理解有帮助。早期的语言模型，可以由左至右地进行训练，或者由右至左地进行训练，但是两者不能同时进行，否则在双向训练过程中会间接地看到答案。
- masked language model：人类理解语言时会同时考虑上下文的语境。BERT 巧妙地借助了完形填空的思想，提出了 masked language model，实现双向 transformer。把输入序列中 15%的词掩盖（Mask），然后通过上下文预测这些被掩盖住的词。
- sentence-level relation：通过预测句子 B 是否句子 A 的下一句话，学习句子间的关系。

BERT 模型将无监督 pre-training+有监督 fine-tuning 这一模式推广到更深层的双向结构中，可以获取长距离语义信息，能够左右上下文语义，可以并行进行。BERT 模型的缺点主要存在于 masked language model 方面：（1）[Mask]标记在实际预测中不会出现，训练时用过多[Mask]影响模型表现；（2）每个 batch 只有 15%的 token 被预测，所以 BERT 收敛得比 left-to-right 模型要慢（它们会预测每个 token）。

- **XLNet**

BERT 模型在预训练阶段引入[Mask]标记，而在下游任务中不存在，破坏了预训练与下游任务的一致性。BERT 在根据句子的其他词预测掩盖掉的单词时，对被掩盖掉的单词做了独立性假设，即假设被掩盖掉的单词之间相互独立。为了解决以上问题，卡内基梅隆大学和谷歌在 2019 年联合推出了 XLNet 模型[166]。XLNet 模型采用自回归语言模型（是一种使用上下文词来预测下一个词的模型）的模式，通过对输入全排列来引入上下文信息（通过双流自注意力机制和 Attention 掩码实现），解决了[Mask]标记带来的负面影响。XLNet 模型使用了最新的 Transformer-XL 模型，直接使用了相对位置编码，并将递归机制整合到全排列设定中[167]。

XLNet 仍然遵循两阶段的过程，第一个阶段是语言模型预训练阶段；第二阶段是任务数据 Fine-tuning 阶段。在 XLNet 模型中，为了实现以下目标：看上去输入句子 X 仍然是自左向右的输入，看到 T_i 单词的上文 Context_before，来预测 T_i 这个单词。但是又希望在 Context_before 里，不仅仅看到上文单词，也能看到 T_i 单词后面的下文 Context_after 里的下文单词。XLNet 把 BERT 的[Mask]的过程搬到 Attention Mask 来完成。这样从输入端看去预训练和 Fine-tuning 是一致的。Attention Mask 的原理是，假设输入的词是 x1–>x2–>x3–>x4，我们在 Attention Mask 中改成随机排列组合的另外一个顺序 x3–>x2–>x4–>x1 了，然后预测 x3 的时候上文为空，预测 x2 的时候上文是 x3，预测 x4 的时候上文是 x3, x2，预测 x1 的时候上文是 x3, x2, x4，这样就达到了预测一个词使用到了上下文的内容。在 Attention Mask 中，输入词的顺序没有变，只是通过 Mask 的操作达到了类似随机排序的效果[167]。

AMiner

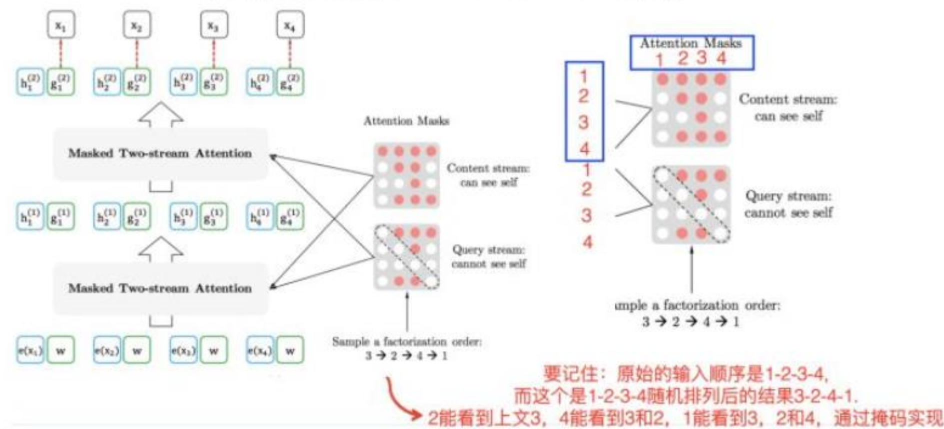


图 43 XLNet 的 Attention Mask 原理机制

XLNet 模型维持了表面看上去的自回归语言模型的从左向右的模式，对于自然语言生成任务，能够在维持表面从左向右的生成过程前提下，模型里隐含了上下文的信息。此外，XLNet 模型引入了 Transformer XL 的机制（相对位置编码以及分段 RNN 机制），对于长文档输入类型的 NLP 任务，也会比 BERT 有明显优势。

AMiner

2.3.3 自然语言生成应用

自然语言生成技术能够大规模的产生个性化内容，加速内容生产，帮助人类洞察数据，让数据更容易理解。由于新闻和报告的内容有比较明显的规则，尤其是金融、体育等领域，因而自然语言生成技术在新闻和报告自动生成方面应比较普遍，通过解读数据，自动输出结论和观点。此外，现在比较流行的聊天机器人、智能音箱、智能客服、电商服务等也使用了自然语言生成技术来回答用户的问题，期待实现更加智能和自然的人机交互。维基百科详细描述了自然语言生成技术在气象报告、文本摘要、写作等方面的应用[168][169]。

大众媒体向来以报导会说笑话的 NLG 系统为重点，不过从商业的角度，最成功的 NLG 应用是从数据库或资料集产生文字摘要的“资料转文本”系统，这些系统通常除了文字生成外，也进行资料分析，特别是几个已经建成的从气象资料产生气象报告文的系统。此类系统最早是加拿大的 FoG，于 1990 年代开始制作英、法语的气象报告，其成功引领了其它研究和商业系统。近年的一个研究显示，使用者有时候偏好电脑产生的气象报告，因为使用的术语比较一致。另一个研究

显示统计的技巧可以用来生成高品质的气象报告。“苏格兰花粉预报系统”就是一个基于版型的简单自然语言生成系统。该系统利用苏格兰几个区域花粉预报的六个输入数目，生成简短的花粉等级文本作为输出。例如使用 2005 年 7 月 1 日的历史资料，该软件的输出：“星期五花粉等级已从昨天的中级到今天的高级，全国大部分地区的数值在 6 到 7。不过北部是中级，数值为 4。”。对比实际的预报（气象员手拟）为：“苏格兰大部分地区花粉量还是很高，为 6 级，东南部更是 7 级。只有北方岛屿及苏格兰东北角情况舒缓，花粉量中等”。

在摘要自动生成方面，自然语言技术的应用也比较成熟。1990 年代自然语言生成开始用于财经和商业摘要。例如 AC 尼尔森研发的 SPOTLIGHT 系统，分析大量零售资料然后自动产生可读的英文。Gartner 已经表示 NLG 将成为 90%的现代商业智能分析平台的标准功能。更近的 NLG 也用于摘要电子医疗记录，已有商业应用问世，研究人员已显示 NLG 摘要的医疗资料可以有效辅助医疗专业人员做决定。

内容生成系统可以辅助真人作者，让写作过程更有效率。例如利用搜索引擎应用界面，基于网络探勘建立的内容生成工具，模拟作者根据各种搜寻结果形成内容，有如剪贴的写作场景。相关度的验证，对于过滤不相关的搜寻结果至为重要，方法基于匹配问题与可能答案的语法树。另一种方法利用真人写的文本建立高层次结构的版型，然后自动写成维基百科新的条目。

自然语言生成是自然语言处理领域的一个重要研究方向，实现文本自动生成也是人工智能走向成熟的一个重要标志。简单来说，我们期待未来有一天计算机能够像人类一样会写作，能够撰写出高质量的自然语言文本。文本自动生成技术极具应用前景。例如，文本自动生成技术可以应用于智能问答与对话、机器翻译等系统，实现更加智能和自然的人机交互。我们也可以通过文本自动生成系统替代编辑实现新闻的自动撰写与发布，最终将有可能颠覆新闻出版行业；该项技术甚至可以用来帮助学者进行学术论文撰写，进而改变科研创作模式。

2.3.4 自然语言生成研究问题与挑战

近年来，随着深度学习技术的发展、大规模数据语料的积累、自然语言处理

AMiner

技术整体发展的进步，以及工业界的现实需求，自然语言生成的相关研究和工程实现受到了诸多关注。越来越多的科技巨头开始看到了这块潜在的“大蛋糕”中蕴藏的价值，他们纷纷通过招兵买马、合作、并购的方式在这方面进行布局。2013年谷歌以超过 3000 万美元收购了新闻阅读应用开发商 Wavii。Wavii 可以通过扫描互联网发现新闻，并给出一句话摘要。国内的科大讯飞在 2016 年底发布了自然语言处理云平台，推出语音合成产品，在中文领域的自然语言处理和语音合成方面有着深厚积累。与此同时，也不断有新兴的科技公司涌现，提出自己的在智能交互、语音识别、机器翻译等方面的解决方案，试图在自然语言生成这片广阔的蓝海上划分自己的领土、树立自己的标杆。但是受限于语言边界模糊、歧义多样、表达不规范、学习语料规模有限、应用场景复杂等多种因素，自然语言生成技术在实际应用中并不成熟，相关研究与实现还存在很多问题与挑战。

自然语言生成技术目前在实际应用中以基于规则的模板生成技术为主，应用在体育、新闻、财经报道等生成文本格式相对固定的领域。虽然能实现文本结构不太复杂的文本自动生成功能，但是结果较为生硬、死板，而且后期扩展对初期规划依赖较大，具体表现在:（1）自动生成文本的格式固定，结构不灵活，不便于修改;（2）生成句子表达不灵活，不符合语言文字表达多样性的特点。

基于深度学习的生成技术不需要人为设计规则模板和语法规划过程，能自动学习到语法和语义规则，可以直接从文本中提取特征。自动生成的文本中，句子的表达更灵活，更符合语言文字表达多样性特点。基于深度学习的方法可以更好的生成创意性、个性化、娱乐性等文本，但是效果比较随机，存在难以控制的风险。目前该项技术研究不是很成熟，存在以下不足：（1）语料库等数据库的构建是基础，目前汉语研究中语料库的构建是一个难点；（2）神经网络结构还不能完全适应文字的处理，虽然目前神经网络结构能实现对文本历史信息的追踪，但仍有诸多缺陷：①容易出现梯度消失和梯度爆炸问题；②不能完全实现对文本信息的控制；③生成的文本仍有语法错误，文本中的语法错误检测和自动纠错机制并没有应用到文本自动生成技术中。现有的基于神经网络的自然语言生成模型对文本自动生成领域的应用研究还远远不够，对新网络结构的探索将是未来研究的一个方向[170]。

自然语言生成技术的另一大瓶颈就是如何精确地表现自然语言的语义，比如在人机交互过程中，首先就要理解用户的意图，而这里“用户的意图”就是语义。目前业界常用的方法有两种：基于知识或语义学规则的语义分析方法和基于统计学的语义分析方法。尽管两类方法都能在一定程度上进行自然语言语义的推导以及信息之间关联的判别，但是基于知识与语义学规则的方法无法覆盖全部语言现象，推理过程复杂，无法处理不确定性事件，规则间的相容性和适用层次范围存在缺陷和限制，知识和语义规则的建立是瓶颈问题；而基于统计学的方法则过多地依赖于大规模语料库的支持，性能依赖语料库的优劣，易受数据稀疏和数据噪声的干扰，正如之前提到，大规模语料库的建立和语料质量的保证仍是瓶颈问题[171]。

2.3.5 自然语言生成未来研究方向

近年来，各种词表、语义语法词典、语料库等数据资源的日益丰富，词语切分、词性标注、句法分析等技术的快速进步，各种新理论、新方法、新模型的出现推动了自然语言处理研究的繁荣，也带动了自然语言生成技术的持续关注与不断发展。自然语言生成因技术难度太大、应用场景太复杂，研究成果还未达到足够的高度。

随着数据多样性的发展、用户需求的增加，传统的基于句法-语义规则的模板生成方法难以满足社会发展的要求。随着深度学习技术的不断发展和成熟，面向自然语言生成的深度学习研究工作已经取得一定成果，但还没有重大突破。在当前已有的深度学习模型研究中，难点是在模型构建过程中参数的优化调整方面，主要有深度网络层数、正则化问题及网络学习速率等，可能的解决方案有采用多核机提升网络训练速度，针对不同应用场合，选择合适的优化算法等。深度学习方法的不可解释和随机性可以通过使用基于模板的生成方法进行补充和完善，将两者方法结合，取长补短，是自然语言生成领域的未来研究方向。

清华大学黄民烈副教授在“知识增强的自然语言生成”报告中指出：自然语言生成中的问题一共分为 4 个维度，第一个是 Information-（Specificity Informativeness Diversity），第二个是 Grammar-（Repetition Fluency），第三个是 Faithfulness-（Fidelity Relevance），第四个是 Coherence-（Consistency Coherence

AMiner

Conflict）。在开放式语言生成（输入信息有限、普遍存在 1 对多映射的语言生成）任务中，多样性、信息量、重复性、一致性、连贯性或逻辑性没有得到很好的控制。此外，他结合学术界现有的工作探讨了一个重要的问题：现有的语言生成模型多大程度拥有常识或世界知识？在常识生成、知识补全、对话生成、故事结局生成、常识故事生成等应用中，知识可以更好地辅助语言理解和语言生成。这些研究尝试表明，知识可以提供信息量、一致性、连贯性和隐式规划，在开放式语言生成中尤其有用。

黄民烈副教授在报告中提到知识图谱可以为自然语言生成带来 Implicit planning、Informativeness、Causal/ temporal dependency、logic coherence 这几方面的提升，在模型中改变原有注意力机制为加入知识图谱编/解码的多源注意力机制，使原本的词向量加上知识图谱中的关系信息；动态图的注意力机制通过加权求和来进行知识图谱关系之间的表达；任务知识图谱在 open-ended 语言生成任务效果提升上起着关键的作用。但是，在没有实际知识以及语义不连贯的情况下进行 open-ended 生成文本或长文本仍然是一个挑战。此外，考虑到知识图谱对自然语言生成技术的提升作用，那么如何构建能够融合符号逻辑和表示学习的大规模高精度的知识图谱，以及如何使用深度神经网络处理自然语言，建立更有效、可解释的语言认知模型也急需探索。

自然语言生成是在自然语言理解的研究发展到一定阶段的必然方向，因此，一方面有关语言理解（如机器翻译）的模型需要取得更好的表现，同时生成模型也需要更灵活。语言的离散性给训练 GAN 造成的困难可以通过结合强化学习缓解，可以帮助生成更真实和灵活的内容。今后自然语言生成的研究发展方向主要是在以下几个方面：第一，将研究重点从规范语法转化为提高语言形式在日常交际中的合理性和有效性，将语言生成变成一种实实在在的交际手段。第二，加大对自然语言生成中信息策划过程的研究，文本信息的策划是研究的重中之重。第三，将句法词汇和信息策划压缩合并也是未来研究的重点。在科学工作者们的通力合作下，自然语言生成在未来肯定会取得更大的突破。

2.4 论文主题分析

为了展示认知图谱领域的研究主题分布情况，旨在为充分了解该领域学术发

AMiner

展进程和学术研究现状提供参考，本章节针对 AMiner 平台上收录的领域学者推荐的 100 篇经典论文（https://www.aminer.cn/topic/5e7d9a78ea0348b7e22edd9a），采用文档主题生成模型（Latent Dirichlet Allocation, LDA），分析这些论文的研究主题分布情况。

LDA[172]假定文本中的每个词由一些混合的主题产生的，每个主题都有一定的权重，即$p(w)=\sum_z p(w|z)p(z)$，其中$p(z)$又是一个 Dirichlet 分布产生。LDA 的贝叶斯网络结构如图 44 所示，图中 K 为主题个数，M 是论文总数，N 是某个论文中单词总数，α和β分别是每个主题下词的多项式分布和每个论文下主题的多项分布的 Dirichlet 先验参数。LDA 模型中有一组隐含变量z，参数求解采用吉布斯采样，构建 Markov 链，逼近目标概率分布。获取参数后可以计算论文的主题关键词，$p(w|d)=\sum_z p(w|z)p(z|d)$。

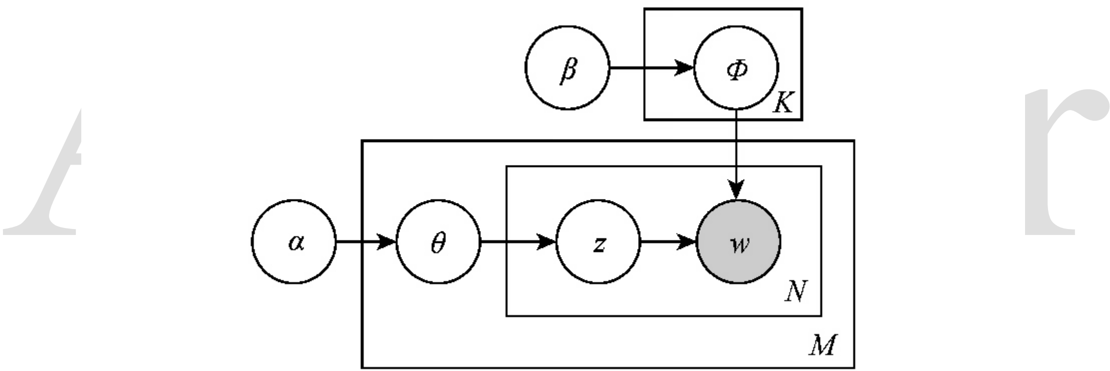


图 44 LDA 结构图

本报告设置主题数 K=5，通过对论文的标题和摘要进行分析，获取了这些论文的研究主题包括：Description Logic（描述逻辑）、Ontology（本体）、Knowledge Reasoning（知识推理）、Text Generation（文本生成）、Knowledge Graph（知识图谱）。表 4 展示了每个主题下具有代表性的 3 篇论文及作者信息。

表 4 认知图谱领域论文主题分布

研究主题	相关论文
Description Logic	标题：A Logical Framework for Default Reasoning 作者：David Poole

	出版信息：Artificial Intelligence, 1988
	标题：Ontology Reasoning in the SHOQ（D） Description Logic 作者：Ian Horrocks, Ulrike Sattler 出版信息：International Joint Conference on Artificial Intelligence, 2001
	标题：Ordering Heuristics for Description Logic Reasoning 作者：Ian Horrocks, Ulrike Sattler 出版信息：International Joint Conference on Artificial Intelligence, 2005
Ontology	标题：A Survey on Ontology Mapping 作者：Namyoun Choi, II-Yeol Song, Hyoil Han 出版信息：ACM SIGMOD Record, 2006
	标题：Ontology Based Context Modeling and Reasoning using OWL 作者：X.H. Wang, D.Q. Zhang, T. Gu ; H.K. Pung 出版信息：Institute of Electrical and Electronics Engineers Annual Conference on Pervasive Computing and Communications Workshops, 2004
	标题：Ontology Learning for the Semantic Web 作者：Alexander D. Maedche 出版信息：IEEE Intelligent systems, 2001
	标题：Knowledge Graph Embedding with Concepts 作者：Niannian Guan, Dandan Song, Lejian Liao 出版信息：Knowledge-Based Systems, 2019
Knowledge Reasoning	标题：From Recognition to Cognition: Visual Commonsense Reasoning 作者：Rowan Zellers, Yonatan Bisk, Ali Farhadi, Yejin Choi 出版信息：Institute of Electrical and Electronics Engineers Annual Conference on Computer Vision and Pattern Recognition, 2019
	标题：Iteratively Learning Embeddings and Rules for Knowledge Graph Reasoning 作者：Wen Zhang, Bibek Paudel, Liang Wang, Jiaoyan Chen, Hai Zhu, Wei Zhang, Abraham Bernstein, Huajun Chen 出版信息：The World Wide Web Conference, 2019

AMiner

	标题：Cognitive Modeling: Knowledge, Reasoning and Planning for Intelligent Characters 作者：John Funge,Xiaoyuan Tu,Demetri Terzopoulos 出版信息：Annual Conference on Computer Graphics and Interactive Techniques, 1999
Text Generation	标题：Topic-Guided Variational Autoencoders for Text Generation 作者：John Funge, Xiaoyuan Tu, Demetri Terzopoulos 出版信息：North American Chapter of the Association for Computational Linguistics, 2019
	标题：Diversity-Promoting GAN: A Cross-Entropy Based Generative Adversarial Network for Diversified Text Generation 作者：Jingjing Xu, Xuancheng Ren, Junyang Lin,Xu Sun 出版信息：Conference on Empirical Methods in Natural Language Processing, 2018
	标题：A Hybrid Convolutional Variational Autoencoder for Text Generation 作者：Stanislau Semeniuta, Aliaksei Severyn, Erhardt Barth 出版信息：Conference on Empirical Methods in Natural Language Processing, 2017
Knowledge Graph	标题：Knowledge Graph Embedding: A Survey of Approaches and Applications 作者：Quan Wang, Zhendong Mao, Bin Wang, Li Guo 出版信息：Institute of Electrical and Electronics Engineers Transactions on Knowledge and Data Engineering Publisher, 2017
	标题：Knowledge Graph Refinement: A Survey of Approaches and Evaluation Methods 作者：Heiko Paulheim 出版信息：Semantic Web, 2017
	标题：Knowledge Graph Completion with Adaptive Sparse Transfer Matrix 作者：Guoliang Ji, Kang Liu, Shizhu He, Jun Zhao

AMiner

	出版信息：the Association for the Advance of Artificial Intelligence Conference, 2016

2.5 经典论文解读

为了帮助读者了解认知图谱相关领域的近期研究成果，本报告整理了 2019 和 2020 年间发表的 10 篇具有代表性的论文，并对论文中的问题、方法、结论等进行了总结。想获知该领域更多论文，可访问 AMiner 平台的"认知图谱"领域经典论文列表，网址如下：

https://www.aminer.cn/topic/5e7d9a78ea0348b7e22edd9a

AMiner 推出了"Topic 必读论文"，收录了计算机领域的 128 个主题的经典论文列表，希望帮助读者快速了解主题概念及代表性经典论文。每个主题中必读论文的选择，一方面是经过大数据分析技术对重要历史节点代表性论文的挖掘，另一方面是热心读者的推荐和整理。读者可以对自己认可的入选论文进行点赞投票，也可以添加心目中的经典必读论文。访问网址：https://www.aminer.cn/topic。

AMiner

1、

论文标题：From Recognition to Cognition: Visual Commonsense Reasoning

论文作者：Rowan Zellers, Yonatan Bisk, Ali Farhadi, Yejin Choi

发表刊物：CVPR, 2019.

论文地址：https://aminer.cn/pub/5c04966a17c44a2c747085f7/from-recognition-to-cognition-visual-commonsense-reasoning

论文解读：针对视觉系统对图像认知和常识推理的问题，该论文介绍了一种将丰富注释转换为具有最小偏差的多项选择问题的新方法，建立了一个由来自 11 万个电影场景的 29 万个选择题组成新的数据集 VCR。为了达到认知水平的图像理解，该论文提出了一个新的推理引擎，称为认知网络识别（R2C），它为基础、语境化和推理建立必要的分层推理模型。实验结果表明，虽然人类发现录像机很容易达到较高准确率（超过 90%的准确率），但最先进的模型只有约 45%的准确率，而该论文提出的 R2C 模型有助于缩小这一差距（约 65%）。

2、

论文标题：Learn to Explain Efficiently via Neural Logic Inductive Learning

论文作者：Yuan Yang, Le Song

发表刊物：ICLR, 2020.

论文地址：https://aminer.cn/pub/5e5e18bf93d709897ce2d1c9/learn-to-explain-

efficiently-via-neural-logic-inductive-learning

论文解读：为提升开发责任机学习系统做出可解释和自解释决策能力，该论文研究了如何在归纳逻辑程序设计（ILP）范围内学习解释的问题。该论文提出了神经逻辑归纳学习（NLIL），这是一个有效的可区分 ILP 框架，能够学习一阶逻辑规则并解释数据中的模式。与最先进的模型相比，NLIL 能够搜索长度长 10 倍的规则，且速度快 3 倍。该研究还发现 NLIL 可以扩展到具有 100 万个实体的大型图像数据集，即视觉基因组。

3、

论文标题：ReClor: A Reading Comprehension Dataset Requiring Logical Reasoning

论文作者：Weihao Yu, Zihang Jiang, Yanfei Dong, Jiashi Feng

发表刊物：ICLR, 2020.

论文地址：https://aminer.cn/pub/5e5e18da93d709897ce35cb4/reclor-a-reading-comprehension-dataset-requiring-logical-reasoning

论文解读：为引入更具挑战性的数据集来推动阅读理解领域朝着更全面的文本推理方向发展，该论文介绍了一种从标准化研究生入学考试中提取的需要逻辑推理的阅读理解数据集。为了综合评价模型的逻辑推理能力，该论文提出了将有偏数据点识别为易集，其余为硬集的方法。实验结果表明，最新的模型在易集上具有很高的精度，能够捕捉到数据集中包含的偏差。然而，它们在硬集上表现较差，性能接近随机猜测，需要更多的研究从本质上提高现有模型的逻辑推理能力。

4、

论文标题：Cognitive Graph for Multi-Hop Reading Comprehension at Scale

论文作者：Ming Ding, Chang Zhou, Qibin Chen, Hongxia Yang, Jie Tang

发表刊物：ACL, 2019.

论文地址：https://aminer.cn/pub/5cf48a37da56291d58299b69/cognitive-graph-for-multi-hop-reading-comprehension-at-scale

论文解读：该论文提出了一个新的 CogQA 框架来解决 web 文档中的多跳问题。在认知科学双过程理论的启发下，该框架通过协调隐式提取模块（System 1）和显式推理模块（System 2），在迭代过程中逐步构建一个认知图谱。该论文基于 BERT 和 graph 神经网络的实现，有效地处理了 HotpotQA-fullwiki 数据集中数百万个用于多跳推理问题的文档，在排行榜上获得了 34.9 的分数，而最佳竞争对手仅为 23.6 分。该框架可以推广到其他认知任务，如会话人工智能和顺序推荐。

AMiner

5、

论文标题： InfoGraph: Unsupervised and Semi-supervised Graph-Level Representation Learning via Mutual Information Maximization

论文作者： Fan-Yun Sun, Jordan Hoffmann, Vikas Verma, Jian Tang

发表刊物： ICLR, 2020.

论文地址： https://aminer.cn/pub/5e5e189a93d709897ce1e760/infograph-unsupervised-and-semi-supervised-graph-level-representation-learning-via-mutual-information

论文解读： 传统的图核方法对于获得图的定长表示是有效的，但是由于手工设计，它们的泛化能力较差。该论文研究了在无监督和半监督两种情况下全图学习的表示，提出了一种新的无监督表示学习方法 InfoGraph。该论文将图级表示和不同尺度子结构（如节点、边、三角形）表示之间的相互信息最大化。此外，还进一步提出了半监督场景 InfoGraph 的一个扩展 InfoGraph*。研究发现受监督的编码器在保留当前受监督任务所偏好的潜在语义空间的同时，能够从未标记的数据中学习。在图形分类和分子性质预测任务上的实验结果表明，InfoGraph 优于最新的基线，InfoGraph*可以实现与最新的半监督模型相竞争的性能。

6、

论文标题： A Comprehensive Survey on Graph Neural Networks

论文作者： Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, Philip S. Yu

发表刊物： arXiv, 2019.

论文地址： https://aminer.cn/pub/5c5ce50d17c44a400fc38c4b/a-comprehensive-survey-on-graph-neural-networks

论文解读： 图形数据的复杂性对现有的机器学习算法提出了严峻的挑战，近年来，关于扩展图形数据深度学习方法的研究层出不穷。该论文综述了图形神经网络在数据挖掘和机器学习领域的应用，提出了将最新的图形神经网络划分为不同的类别的新分类法。以图形卷积网络为重点，该论文回顾了最近发展起来的替代体系结构；这些学习范式包括图形注意网络、图形自动编码器、图形生成网络和图形时空网络。该论文进一步讨论了图神经网络在各个领域的应用，总结了现有算法在不同学习任务上的开源代码和基准。最后，该论文提出了这一领域的潜在研究方向。

7、

AMiner

论文标题： Explainable Reasoning over Knowledge Graphs for Recommendation

论文作者： Xiang Wang, Dingxian Wang, Canran Xu, Xiangnan He, Yixin Cao, Tat-Seng Chua

发表刊物： AAAI, 2019.

论文地址： https://aminer.cn/pub/5c04967517c44a2c74708b3b/explainable-reasoning-over-knowledge-graphs-for-recommendation

论文解读： 近年来，将知识图谱引入推荐系统引起了越来越多的关注。然而现有的工作并没有充分探索这种连接性来推断用户偏好，特别是在建模路径内的顺序依赖关系和整体语义方面。基于该问题，该论文提出了一种新的知识感知路径递归网络（KPRN）模型，利用知识图谱进行推荐。KPRN 可以通过组合实体和关系的语义来生成路径表示。通过利用路径中的顺序依赖关系对路径进行有效推理，以推断用户项交互的基本原理。该论文还设计了一个新的加权池操作来区分不同路径在连接用户和项目时的优势，使模型具有一定的可解释性。该论文在两个有关电影和音乐的数据集上进行了广泛的实验，展示了与最先进的解决方案-协作知识库嵌入和神经分解机相比，KPRN 模型具有显著的改进。

8、

论文标题： Topic-Guided Variational Autoencoders for Text Generation

论文作者： Wenlin Wang, Zhe Gan, Hongteng Xu, Ruiyi Zhang, Guoyin Wang, Dinghan Shen, Changyou Chen, Lawrence Carin

发表刊物： ACL, 2019.

论文地址： https://aminer.cn/pub/5cede0efda562983788ce72b/topic-guided-variational-autoencoders-for-text-generation

论文解读： 基于 VAE 的文本生成模型采用简单的高斯先验，不能表示句子之间的语义结构，还存在"后塌陷"的问题。通过削弱对解码器的条件依赖性的解决方法可能导致无法生成高质量的连续句子的问题。该论文提出了一个主题导向的变分自动编码器（TGVAE）文本生成模型。与现有的基于变分自编码（VAE）的方法不同，该方法假定隐码具有简单的高斯先验，模型将先验指定为一个由神经主题模块参数化的高斯混合模型（GMM）。每一个混合成分对应一个潜在的主题，提供在主题下生成句子的指导。模型中的神经主题模块和基于 VAE 的神经序列模块是联合学习的。在模型推理过程中，利用一系列可逆 Householder 变换赋予隐码的近似后验以高度的灵活性。实验结果表明，TGVAE 在无条件文本生成和条件文本生成方面都优于其他方法，可以生成具有不同主题的语义意义的句

AMiner

子。

9、

论文标题： Knowledge graph embedding with concepts

论文作者： Niannian Guan, Dandan Song, Lejian Liao

发表刊物： Knowledge Based Systerm, 2019.

论文地址： https://aminer.cn/pub/5c3f1e5adf5b8c0b3cce7166/knowledge-graph-embedding-with-concepts

论文解读： 现有的知识图谱嵌入模型主要关注实体-关系-实体三元组，或者与文本语料库交互。然而，三元组的信息量较小，域内文本语料库并不总是可用的，使得嵌入结果偏离了实际意义。该论文基于概念图中实体的常识概念信息，提出了概念嵌入知识图谱（KEC）模型，将实体和实体概念联合嵌入到语义空间中。知识图谱中的事实三元组根据概念图中实体的常识概念信息进行调整。该模型不仅关注实体之间的相关性，而且还关注它们的概念，因此该模型提供了精确的语义嵌入。该论文从知识图谱完成和实体分类两方面对此方法进行了评价。实验结果表明，该模型在两个任务上都优于其他基线。

10、

论文标题： Towards Knowledge-Based Personalized Product Description Generation in E-commerce

论文作者： Qibin Chen, Junyang Lin, Yichang Zhang, Hongxia Yang, Jingren Zhou, Jie Tang

发表刊物： KDD, 2019.

论文地址： https://aminer.cn/pub/5cede0edda562983788cb8c2/towards-knowledge-based-personalized-product-description-generation-in-e-commerce

论文解读： 高质量的产品描述对于在电子商务平台上提供具有竞争力的客户体验至关重要。但自动生成产品描述现有的方法主要是使用模板或统计方法，其性能相当有限。该论文将神经网络和知识库相结合，探索出一种新的个性化产品描述生成方法，提出了一个基于知识的电子商务背景下的产品描述生成模型（KOBE），将编码器-解码器框架“Transformer”扩展为使用自关注的序列建模公式，在文本生成过程中考虑了多种重要因素，包括产品方面、用户类别和知识库等。在实际数据集上的实验表明，该方法在各种度量上都优于基线，与 BLEU 的水平相比提高了 9.7%。目前，该框架已经在中国最大的在线电子商务平台——淘宝上得到应用。

AMiner

2.6 技术情报挖掘

图 45 展示了本报告在前述内容中介绍的认知图谱相关技术。本章节分别选择了“知识图谱—知识图谱构建—信息抽取—实体抽取”、“认知推理—基于分布式表示推理的方法—基于距离模型的表示推理”、“逻辑表达—基于神经网络的自然语言生成—Transformer 模型”等三条线路，来深入挖掘这些技术发展历程中的代表性人物和事件。

- **知识图谱—知识图谱构建—信息抽取—实体抽取**

2012 年，Google 的 Singhal 等人在博客 *Introducing the Knowledge Graph: things, not strings*[173]，中介绍了知识图谱（Knowledge Graph）的概念。知识图谱以语义网络的结构化方式描述客观世界中概念、实体、事件以及它们之间的关系，相对于传统的本体和语义网络而言，实体覆盖率更高，语义关系也更加复杂而全面。Google 使用语义检索从多种数据来源收集信息构建知识图谱，以提高 Google 搜索的质量。知识图谱在 Google 搜索引擎上的成功应用，引起了学术界和工业界的极大研究热情。目前，知识图谱产品不仅有英文、中文、西班牙文等多个语言版本，还延伸到多个领域，如医疗、金融、教育等。

1997 年，纽约大学的 Ralph Grishman 发表了名为 *Information extraction: Techniques and challenges* 的论文[174]，提出了信息抽取（Information Extraction）的概念。信息抽取是从自然语言文本中抽取指定类型的实体（Entity）、关系（Relation）、事件（Event）等事实信息，并形成结构化数据输出的文本处理技术。

AMiner

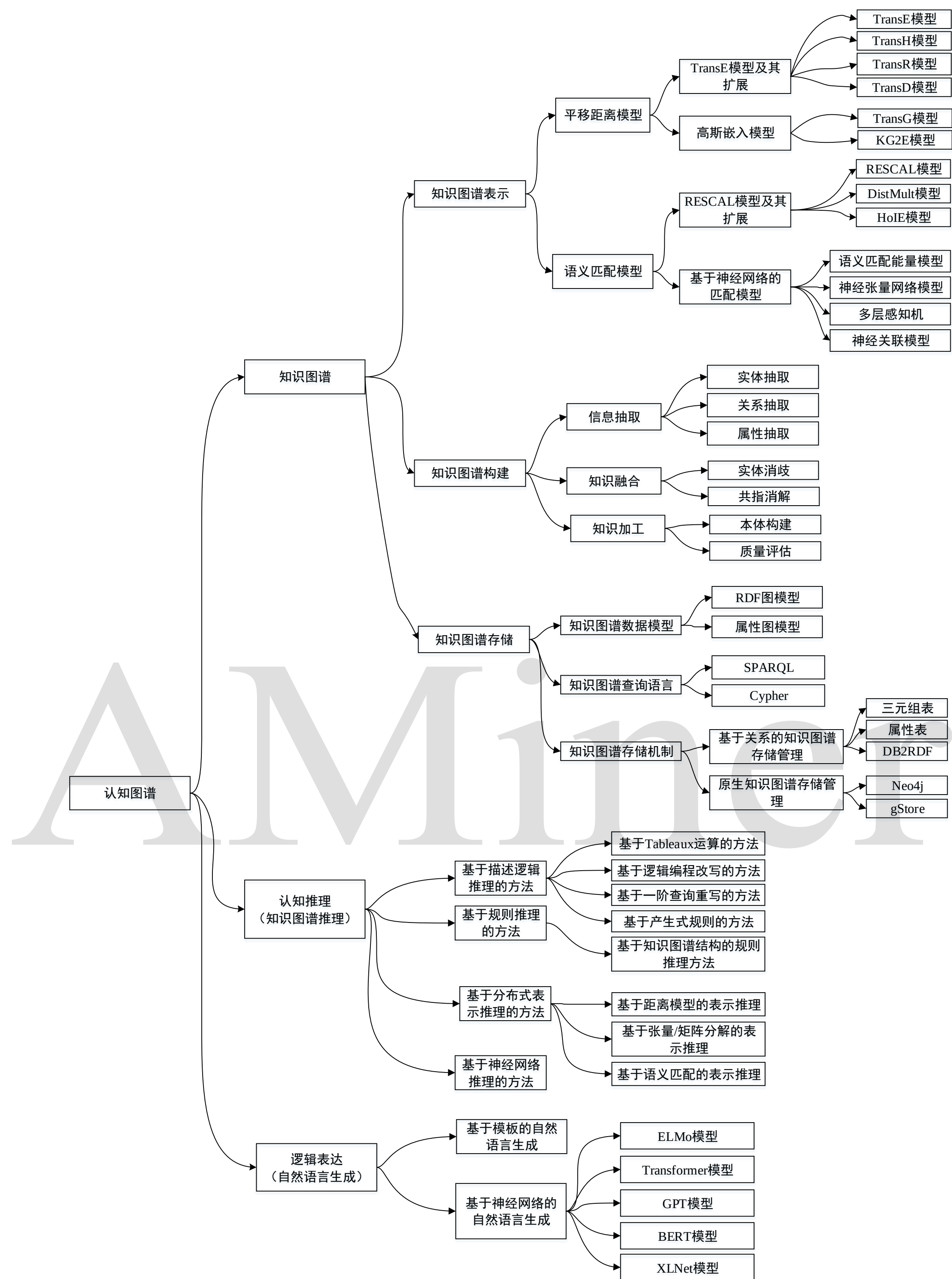


图 45 认知图谱相关技术概览

2008 年，Etzioni 等学者发表了一篇关于开放式信息抽取的经典论文 *Open information extraction from the web*[175]，与传统信息抽取技术相比，开放式信息抽取技术能从海量、冗余、异构、不规范、含有大量噪声的网页中大规模地抽取开

放类别的实体、关系、事件等多层次语义单元信息，并形成结构化数据格式输出，目前开放式信息抽取仍然是主要研究方向。

2007 年，加拿大国家研究理事会的 Nadeau 和纽约大学的 Sekine 联合发表了论文 *A survey of named entity recognition and classification*[176]，对命名实体识别和分类的研究现状进行了详细阐述和总结。命名实体识别（Named Entity Recognition，简称 NER）的任务是识别出待处理文本中三大类（实体类、时间类和数字类）、七小类（人名、机构名、地名、时间、日期、货币和百分比）命名实体。NER 是信息抽取的一项基础性关键任务，同时 NER 也是关系抽取、事件抽取、知识图谱、机器翻译、问答系统等诸多自然语言处理任务的基础。为了应对大规模网络文本的实体抽取需求，部分研究者在 NER 基础上，研究开放式实体抽取技术。开放式实体抽取（Open Entity Extraction）关注的是从海量、冗余、不规范的网络数据源上抽取出符合某个语义类的实体列表。在互联网应用领域，开放式实体抽取技术对于知识库构建、网络内容管理、语义搜索、问答系统等都具有重要应用价值。

AMiner

- **认知推理—基于分布式表示推理的方法—基于距离模型的表示推理**

认知推理旨在结合人脑的推理过程，协同结构化的推理过程和非结构化的语义理解，帮助机器获得多模感知能力，赋能海量任务。知识图谱为认知推理提供了强有力的知识支撑，认知推理的底层实现是知识推理，面向知识图谱的知识推理旨在根据现有知识图谱中的实体、关系等信息，通过推理机制，预测实体之间的关系，以此完善知识图谱，也被称为知识图谱补全。

基于距离模型的表示推理方法是根据转移假设（将关系看成实体间的转移）设计得分函数，度量事实三元组的有效性分值，分值越高表示该元组的有效性越高（即正例元组的分值高，而负例元组的分值低）。2013 年，法国贡比涅工业大学的 Antoine Bordes 等人发表的 *Translating embeddings for modeling multi-relational data*[178]论文中提出的 TransE 是第一个基于距离的表示模型。该模型的主要思想是如果三元组（头实体，关系，尾实体）成立，头实体向量 h 与关系向量 r 的和与尾实体向量 t 相近，否则远离。TransE 模型的提出，促进了一系列知识图谱表示模型和基于分布式表示推理模型的研究。

基于分布式表示推理的方法是指将知识图谱上的事实三元组中实体和关系等元素表示为低维向量形式，基于向量表示形式执行推理预测操作。2015 年，清华大学 Yankai Lin 等人在人工智能顶级会议 AAAI 上发表的论文 *Learning entity and relation embeddings for knowledge graph completion*[177]，从分布式表示推理的角度，研究知识图谱补全方法。该论文提出的 TransR 模型，首先在单独的实体空间和关系空间建立实体和关系的嵌入表示；然后将实体从实体空间投影到对应的关系空间，并在投影实体之间建立平移。TransR 模型是一种基于距离模型的表示推理方法，可以解决早期方法在处理一对多、多对一以及多对多等复杂关系时存在的一些不足和缺陷。

- **逻辑表达—基于神经网络的自然语言生成—Transformer 模型**

逻辑表达是一门研究思维形式结构及其基本规律的科学。自然语言生成是研究使计算机具有人一样的表达和写作的功能，为了自动生成一段高质量的自然语言文本，必须具备良好的逻辑表达能力。

1948 年，美国数学家 Claude Elwood Shannon 把离散马尔科夫过程的概率模型应用于描述语言的自动机，打开了自然语言生成的研究开端。1996 年，阿伯丁大学的 Ehud Reiter 发表的论文 *Building Natural Language Generation Systems*[179]，提出了自然语言生成的三个任务，分别是内容确定和文本规划（Content Determination and Text Planning）、语句规划（Sentence Planning）和实现（Realization）。自然语言生成的研究目标是让计算机具备像人类一样的表达和写作功能，即输入一些关键信息，基于语言信息处理的计算机模型，自动生成高质量、人类可读的自然语言文本。自然语言生成可用于自动化生成长篇内容报告，以及为 Web 或移动应用程序生成自定义内容，也可以用于在交互式对话中生成简短的文本。Gartner 在发布的“商业智能与分析平台魔力象限”报告中预测：到 2020 年，自然语言生成和人工智能将成为 90％现代商业智能平台的标配。

神经网络启发于人类大脑的工作原理，通过对输入和输出数据之间的非线性关系进行建模，为计算提供了一种新的方法，被广泛用于自然语言生成任务中。2019 年，美国微软研究院的 Li Dong 等人在人工智能顶级会议 NeurIPS 发表的论文 *Unified Language Model Pre-training for Natural Language Understanding and*

AMiner

Generation[180]中提出的 UNILM 模型，是一种基于神经网络的统一预训练语言模型，在很多自然语言生成、自然语言理解等任务中有非常好的效果。UNILM 模型是一种多层 Transformer 网络，使用了三种类型的语言建模任务进行预训练：单向模型、双向模型、序列到序列预测模型。其统一建模的实现使用了共享的 Transformer 网络，并且还使用了特定的自注意掩码来控制预测条件所处的上下文。该模型在 2019 年世界互联网大会中荣获“世界互联网领先科技成果奖”。

Transformer 模型是 Google 的 Vaswani A 等人在 2017 年发表的文章“*Attention Is All You Need*”[181]中提出的。该模型摒弃了传统的 RNN/CNN 模型，主要基于 self-attention 机制设计，是第一个用纯 Attention 搭建的模型，将 Attention 思想发挥到极致，在计算速度上得到很大提升。Transformer 模型广泛应用于机器翻译、问答系统、文本生成、语音识别等自然语言处理领域。

AMiner

3　人才篇



3 人才篇

依托清华大学自主研发的"科技情报大数据挖掘与服务平台[1]"(简称 AMiner),采用大数据分析与挖掘技术,结合文献计量学方法和引文分析方法,基于附录 1 和附录 2 列出的认知图谱相关领域的关键词列表、代表性期刊和会议列表,获取这些期刊和会议上在 1950 年—2020 年间收录的相关论文数据(共计 60,735 篇,论文引用量是 2,381,315),通过分析和挖掘这些论文作者的信息,获取了 91,435 位研究学者。以这些学者以及他们发表的论文作为底层数据,对国家和机构的学术水平、学者合作和流动情况进行一个整体分析,并利用 AMiner 平台的学者画像功能,展示了国内外一些代表性学者的详细信息。通过以上数据的挖掘与分析,旨在为国家和机构在认知图谱领域的学科建设和布局、人才政策制定等提供数据参考。

3.1 学者情况概览

AMiner

本章节通过深度分析以上论文和学者数据,从分布地图、学术水平、国际合作、学者流动等多种维度,详细介绍了全球和国内学者在认知图谱相关领域的发展情况。

3.1.1 学者分布地图

人才的聚集可以推动一个城市相关产业的快速发展,人才地图的分布可以直观展示人才的地区分布,有利于调查和分析各地域人才竞争力的现况,对人才调查和引进尤为重要。

本节参考 h-index 作为筛选条件,选择 TOP 5000 全球学者,以地图形式展示和分析他们的地区分布。每个人才分布地图,均是根据科技人才的当前单位地理位置进行绘制,其中颜色越红、圆圈越大,表示人才越集中。

(1) 全球学者分布地图

图 46 展示了认知图谱相关领域的研究学者(h-index TOP 5000)在全球范围

[1] 简称 AMiner(https://aminer.cn/),2006 年上线,已建立起全领域学术论文文献及专家学者库,收录论文\专利文献超过 3 亿篇,学者信息 1.36 亿份。

的分布情况。从中可以看出，这些学者主要集中在以美国为首的北美洲、以英国为首的欧洲，和以中国为首的亚洲。表 5 展示了这些学者数量排名前 10 的国家。美国的高水平学者数量在全球位居首位（学者数量占比约 18.5%），其次是中国（学者数量占比约 13.1%），第三是英国（学者数量占比约 4.6%）。中美两国的高水平学者数量远高于其他国家，在认知图谱领域具有较强的发展潜力。中国虽然仅次于美国，但是高水平学者数量约是美国的一半，相差较大。中国有必要采取相关措施，包括人才培养、引进和激励等策略，来建设高水平人才队伍。



图 46 全球学者分布地图

表 5 h-index TOP 5000 全球学者的国家统计

排名	国家名称	学者数量	排名	国家名称	学者数量
1	美国	923	6	法国	153
2	中国	653	7	加拿大	135
3	英国	229	8	西班牙	133
4	德国	207	9	日本	123
5	意大利	203	10	新西兰	88

（2） 中国学者分布地图

图 47 和表 6 展示了认知图谱相关领域的研究学者（h-index TOP 5000）在中国范围的分布情况。从地图中可以看出，这些高水平的研究学者主要分布在中国的经济和政治发达地区，包括以北京为首的京津冀地区、以上海为首的长江三角洲地区、珠江三角洲、香港、台湾等地区，分布在中部、西部和东北地区的学者

较少。人才的地区分布会影响科研资源的配置，进一步加快人才向发达城市流动的速度，形成资源与人才的循环，增大城市之间的发展差距。政府应该加大经济弱势地区的科研投入和立项力度，通过科技政策协调各个地区的科研资源配置，改善人才地区分布集中的情况。同时应该加强地区之间的人才合作交流，使领域技术在各地区平衡发展。



图 47 中国学者分布地图

表 6 h-index TOP 5000 中国学者的省市统计

排名	省市名称	学者数量	排名	省市名称	学者数量
1	北京	129	6	香港	27
2	上海	45	7	广东	26
3	江苏	36	8	台湾	25
4	辽宁	36	9	浙江	21
5	湖北	27	10	四川	15

3.1.2 学术水平分析

本节通过分析指定期刊和会议上在 1950 年—2020 年间收录的相关论文数据（共计 60,735 篇）和 91,435 位研究学者，统计分析了全球各个国家和机构的学

术水平情况。

（1） 国家学术水平分析

表 7 显示了论文总被引频次排名前 10 的国家的论文发表量、学者数量等统计情况。这些国家主要分布在欧洲（6 个）、北美洲（2 个）、亚洲（1 个）、大洋洲（1 个），由此看出欧洲国家在认知图谱相关领域发表论文的影响力较高。论文总被引频次最高的国家是美国，其次是英国、德国，中国排名第 4。论文发表量和学者数量最高的是美国，第 2 是中国，第 3 和第 4 分别是英国、德国。结合前述的 h-index 前 5000 学者在各个国家的分布情况，中国排名第 2，而英国和德国分别排名第 3 和第 4。以上数据从侧面反映出，中国学者的论文发表量较高，但是论文被引频次偏低。中国相关部门需要重视并采取相关措施，让中国学者在重视论文数量的同时，也要提升论文质量。

表 7 论文总被引频次排名前 10 的国家

国家名称	论文总被引频次	论文发表量	学者数量
美国	1206338	16046	26824
英国	211080	3878	6762
德国	187125	3672	6278
中国	185899	8222	15429
加拿大	127831	2138	3818
意大利	83661	2255	3778
法国	69682	2156	3948
荷兰	61346	1262	2201
西班牙	60079	1719	1828
澳大利亚	50532	1406	2569

（2） 机构学术水平分析

表 8 显示了论文总被引频次排名前 10 的全球各个机构的论文发表量、学者数量等统计情况。这些机构主要分布在美国（5 个）、英国（2 个）、中国（2 个）、德国（1 个）。中国两所机构均在北京，分别排名第 9 和第 10。按照论文发表量排序，中国科学院排在第 4、清华大学排在第 6。按照学者数量排序，清华大学排在第 5、中国科学院排在第 9。清华大学和中国科学院的论文发表量与谷歌公司相差不多，但是论文总被引频次约是谷歌的十分之一。与前述章节分析各个国

AMiner

家的学术水平相似，中国机构的论文发表量和学者数量较高，但是论文总被引频次较低，中国的研究机构有必要提高发表论文后续影响力的重视。

表 8 论文总被引频次排名前 10 的全球机构

机构名称	论文总被引频次	论文发表量	学者数量
谷歌 Google	105604	454	799
斯坦福大学 Stanford University	77836	537	819
卡内基梅隆大学 Carnegie Mellon University	73756	841	1371
微软 Microsoft	52866	579	931
爱丁堡大学 University of Edinburgh	40653	417	614
华盛顿大学 University of Washington	35887	317	534
牛津大学 University of Oxford	30926	263	289
亚琛工业大学 Rwth Aachen University	21414	239	318
清华大学 Tsinghua University	13736	443	774
中国科学院 Chinese Academy of Sciences	10887	457	505

AMiner

表 9 显示了论文总被引频次排名前 10 的中国机构的论文发表量、学者数量等统计情况。这些机构主要分布在北京（3 个）、浙江（1 个）、上海（2 个）、江苏（2 个）、哈尔滨（1 个）、台湾（1 个）。论文总被引频次最高的清华大学，它的学者数量也是最高的，论文发表量排名第 2。论文发表量最高的是中国科学院，但是其学者数量排名第 3，位于北京大学后面。此外，这 10 所机构中存在 7 所是中国 C9 高校联盟成员，分别是清华大学、北京大学、浙江大学、上海交通大学、复旦大学、哈尔滨工业大学、南京大学。从以上数据分析结果可以看出，中国技术领先的科研机构主要分布在经济和政治发达地区，科技创新是经济发展的关键因素，科研弱势容易进一步拉大欠发达地区和发达地区的发展差距。

表 9 论文总被引频次排名前 10 的中国机构

机构名称	论文总被引频次	论文发表量	学者数量
清华大学 Tsinghua University	13736	443	774
中国科学院 Chinese Academy of Sciences	10887	457	505
北京大学Peking University	8624	381	572
浙江大学Zhejiang University	4372	198	382
上海交通大学 Shanghai Jiao Tong University	3889	248	318
南京大学 Nanjing University	3082	113	206
哈尔滨工业大学 Harbin Institute of Technology	2792	211	290
复旦大学Fudan University	2161	154	154
国立台湾大学 National Taiwan University	2559	75	136
东南大学 Southeast University	2115	122	166

AMiner

3.1.3 国际合作分析

（1） 全球国家合作分析

通过统计全球不同国家在认知图谱相关领域的合作论文发表情况，表 10 展示了合作论文数量排名前 10 的国家之间合作情况。从中可以看出，大部分国家倾向与美国、英国合作，与美国合作的国家数量为 8 个，与英国合作的国家数量为 2 个。国家之间合作论文数量排名第一的是中国与美国之间的合作，比排在第二位的英国与美国的合作论文数量多出 1 倍。但是中美两国合作论文的总被引频次稍低于英国与美国的合作论文，排在第二位。与英国具有论文合作关系的其他国家来看，德国第一，其次是中国，这两个国家与英国的论文合作数量相差不多。从数据中可以看出，美国和英国是论文合作国家数量较多的国家，也从侧面反应了这两个国家在该领域的技术较为领先。

表 10 合作论文数量排名前 10 的国家列表

国家 1	国家 2	合作论文数量	论文总被引频次
中国	美国	1005	51149
英国	美国	517	52264
德国	美国	447	40249
加拿大	美国	377	33974
日本	美国	261	13215
德国	英国	227	22186
法国	美国	223	15058
中国	英国	206	8485
印度	美国	202	9522
意大利	美国	178	11036

（2） 中国与其他国家合作分析

图 48 展示了中国与其他国家在认知图谱相关领域的论文合作情况。从中可以看出，中国与美国合作的论文数量最多，其次是英国、新加坡、澳大利亚、加拿大。中美两国合作论文的数量约占中国与其他国家合作论文数量的 47.9%，远高于中国与其他国家合作的论文。中国论文发表量是 8205 篇，中美两国合作论文数量约占比 12.2%。由此可以看出，中国学者倾向与美国合作，并且合作论文数量在中国论文发表量中占比较重，但是现在中美贸易竞争日趋激烈，中国有必要加强与其他国家（比如英国）的合作力度。

AMiner

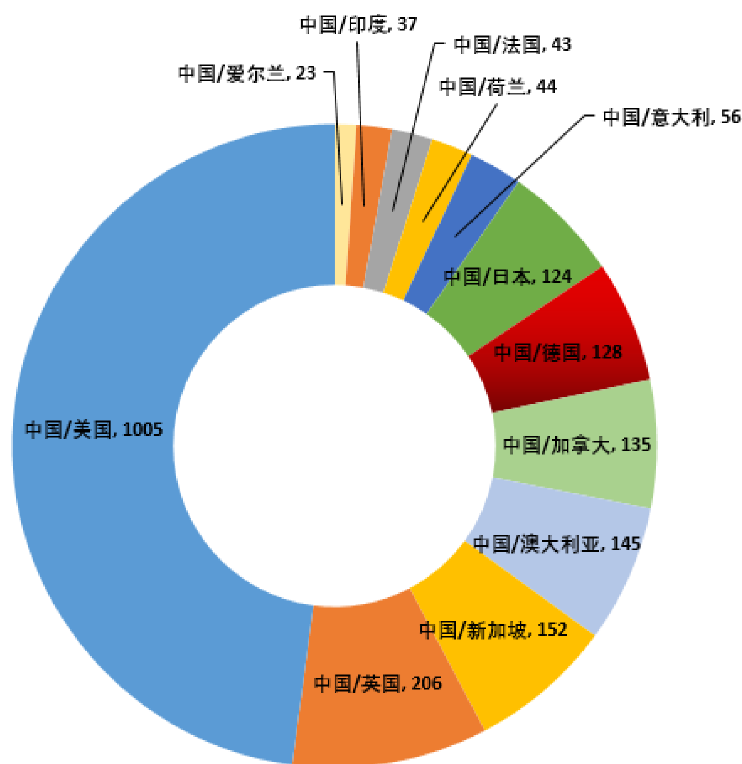


图 48 中国与其他国家的论文合作情况

AMiner

3.1.4 学者流动情况

人才聚集可以推动一个国家或地区的快速发展，而人才的流动是实现人才聚集的重要途径，人才迁徙是科技人才竞争日益激烈时代的鲜明特色。基于学者所属单位地理位置的变化进行统计，呈现全球各个国家和中国各个省市在认知图谱相关领域学者的流入、流出、流动差值等情况，有助于明确各个地区对人才培养、支持和引进的重视程度。

（1） 全球学者流动分析

将学者流入和流出数量之和作为筛选条件，统计了认知图谱相关领域的全球学者在 2009 年—2020 年间的流动数量排名前 10 的国家，如图 49 所示。其中，美国、加拿大、法国的学者流入数量大于流出数量，而中国、英国、德国、意大利、印度、日本、西班牙的学者流出数量大于流入数量。无论是学者流入还是学者流出，美国排名第一，其次是中国。此外，分析中国学者在 1950 年到 2009 年的流动数据，发现中国学者流入的数量为 35，流出数量为 37。由此看出，中国近十年来的学者流动较为频繁，这与中国制定的出国培养、人才引进等相关人才

政策不无关系。随着经济全球化深入发展，科技人才的跨国流动更加频繁，各国对科技人才的争夺更加激烈。人才是科技创新最关键的因素，是提升国家综合国力的基础。中国制定的一系列科技人才培养、引进和激励等政策和措施已取得一定效果，但是有必要对现有科技人才政策进行查缺补漏，建设一支更富有国际竞争力的高水平科技人才队伍。

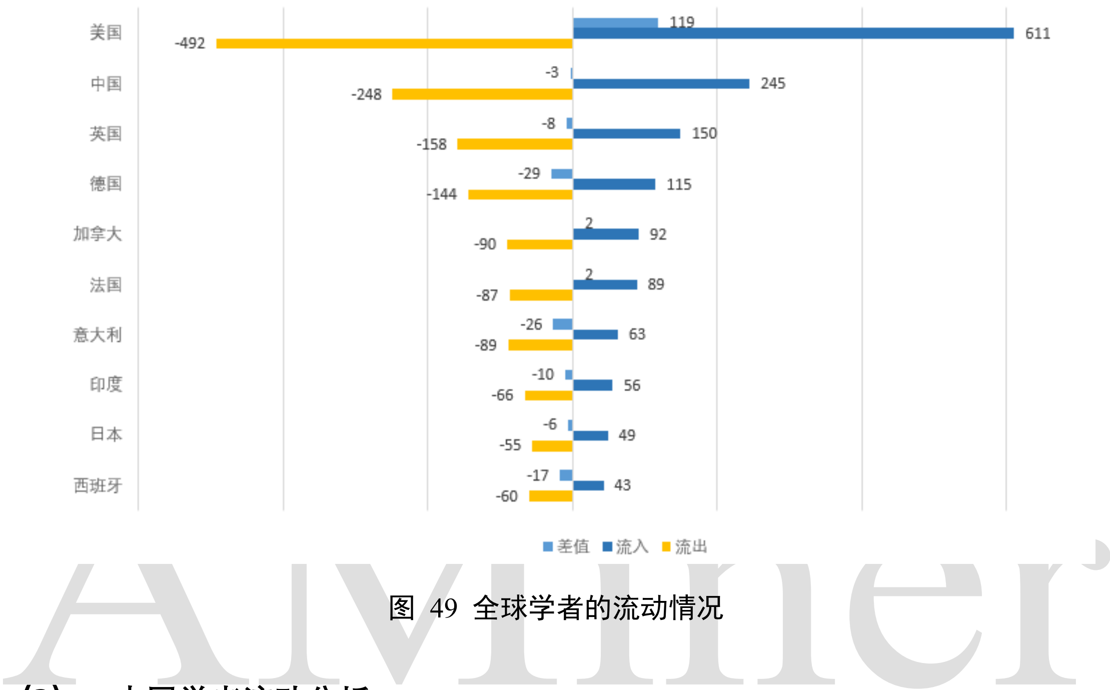


图 49 全球学者的流动情况

AMiner

（2） 中国学者流动分析

本节统计了认知图谱相关领域的中国学者于 2009 年—2020 年间在中国不同省市的流动情况，并根据学者流入和流出数量之和，展示了排名前 10 的中国省市，如图 50 所示。北京是学者流入和流出数量都很高的城市，比其他城市高出较多，这可能与北京的高校和科研院所数量较多有一定关系。此外，北京、香港、上海、哈尔滨的学者流出数量高于流入数量，而武汉、南京、西安、杭州、广州、合肥的学者流入数量高于流出数量。北京、香港、上海的学者流失与这三个城市的高校数量较多、供求关系不平衡、人才竞争力和生活压力较大不无关联，而哈尔滨的地域环境造成了一定的学者流失。针对学者净流入数量为正值的这几个城市，通过调研相关资料，可以发现这些地方在积极推行人才引进政策，包括落户、房补、薪酬等方面均制定了相应的优惠措施。

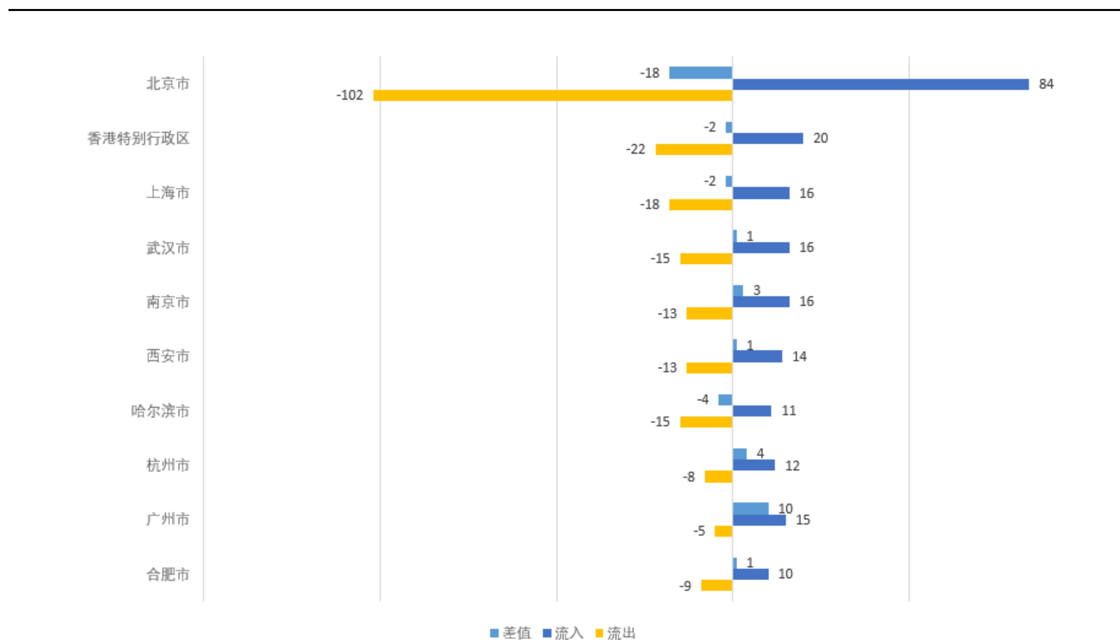


图 50 中国学者的流动情况

3.2 代表性学者画像

本章节依托 AMiner 平台，参考 h-index 指标，针对认知图谱相关领域，筛选了国外和国内各 10 位代表性学者进行学者画像展示和简要介绍（排名不分先后）。限于报告篇幅，本报告不能对所有学者逐一罗列，如有疏漏，还请与 AMiner 编者联系，或者登录 https://www.aminer.cn/获取学者的更多资料。

“学者画像”是 AMiner 平台的核心服务功能之一，具体示例如图 51 所示。学者画像的特色在于除了提供专家学者如姓名、单位、地址、联系方式、个人简介、教育经历等个人基本信息之外，还利用团队多年的命名排歧相关技术基础，建立了较为完全的学者-论文映射关系，提供学者学术评价、研究兴趣发展趋势分析、学者合作者关系网络等分析挖掘信息，同时支持用户交互，通过众包方式丰富专家学者的相关信息。学者的学术指标数值是根据 AMiner 平台的不完全统计，如有论文遗漏请联系编者。

表 11 学术指标说明

学术指标名称	学术指标解释
Papers	一个学者在其学术生涯中发表的所有论文。
Citation	一个学者的所有论文的总被引频次。

H-Index	一个学者的所有论文列表中，至少有 H 篇论文，每篇论文的被引频次均在 H 值以上。
G-Index	一个学者的所有论文列表中，最高引用的 G 篇论文平均至少被引用 G 次。
Sociability	表示领域社交性，通过利用学者和每位合作者之间合著论文数量来度量。
Diversity	表示学科多样性，通过考虑学者发表的所有论文的研究主题数量来度量。
Activity	表示学术活跃度，通过综合考虑论文及其发表期刊或会议的影响力来度量。



图 51 AMiner 平台学者画像示例

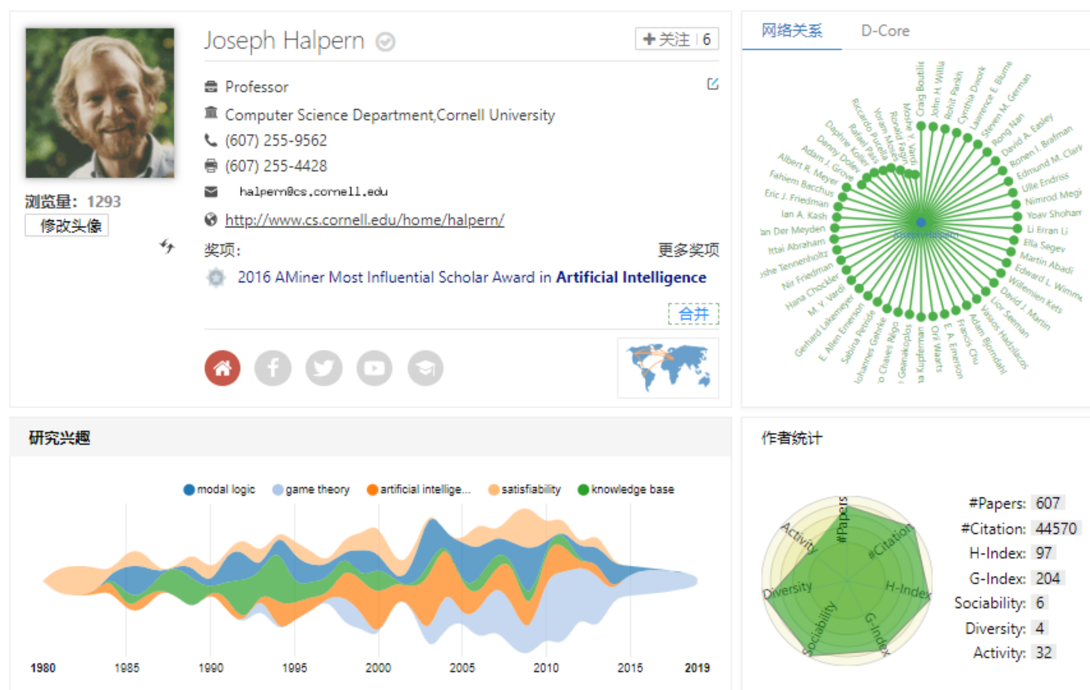
3.2.1 国外代表性学者

● **Amit P. Sheth**



Amit P. Sheth（h-index：103），是美国南卡罗来纳大学人工智能研究所的创办者，在俄亥俄州立大学获得硕士和博士学位。Sheth 教授的当前兴趣包括人工智能（主要关注：知识图谱、自然语言处理、深度学习、知识增强学习、健康和教育方面的 AI 聊天机器人）、语义 Web、大数据、增强个性化健康，以及大数据在健康和生命科学、社会公益、灾害管理等领域的应用。

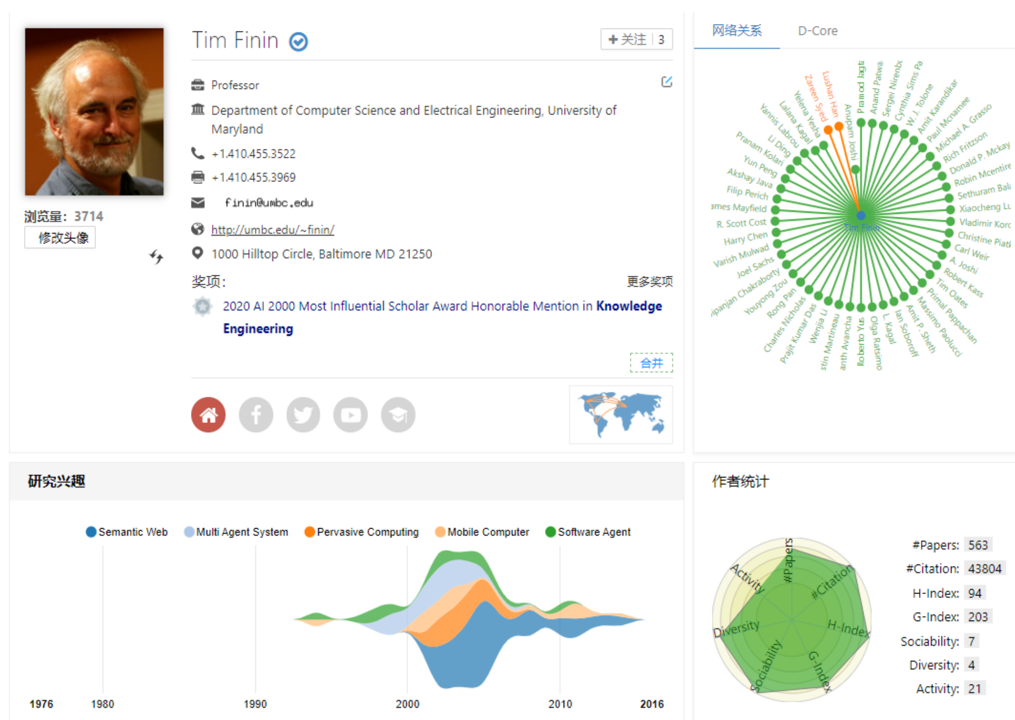
● Joseph Halpern



AMiner

Joseph Halpern（h-index：97），是康奈尔大学计算机科学系教授，曾经担任耶路撒冷希伯来大学客座教授、荷兰阿姆斯特丹大学客座教授、康奈尔大学认知研究项目主任、斯坦福大学咨询教授、多伦多大学客座教授、IBM 阿尔马登研究中心数学和计算机科学系经理、斯坦福大学咨询副教授及咨询助理教授、IBM Research 研究人员、哈佛大学和麻省理工学院博士后研究员、哈佛大学数学教学研究员等。Halpern 教授的主要研究方向是博弈论与决策理论与计算机科学的交叉、知识与不确定性的推理和因果关系。

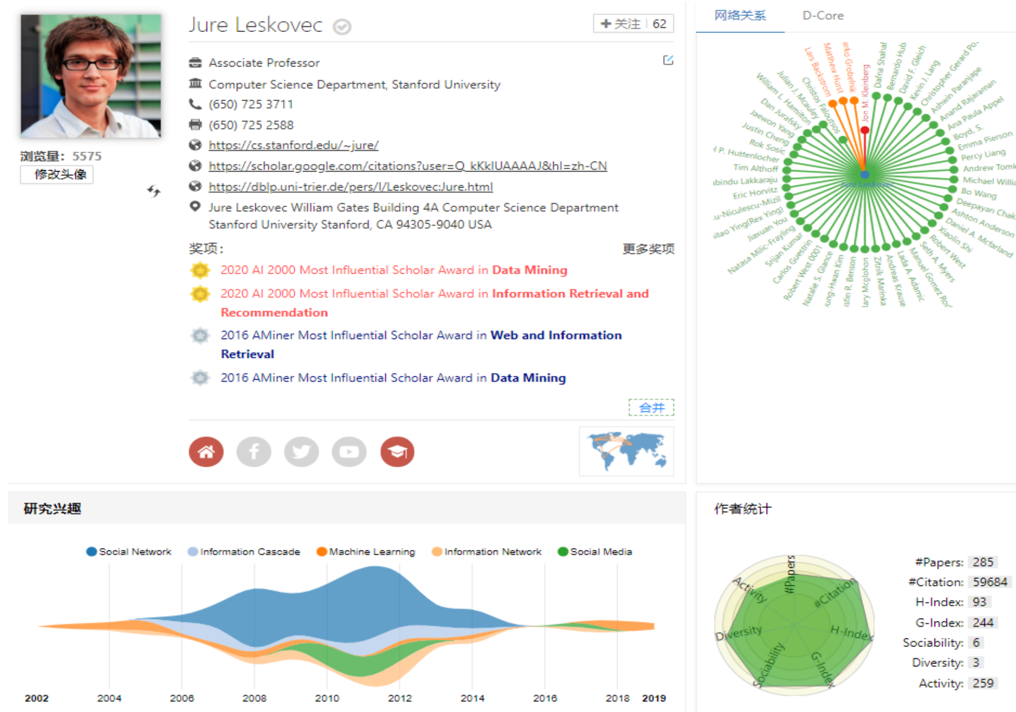
● **Tim Finin**



AMiner

Tim Finin（h-index：94），是 ACM 研究员，AAAI 研究员，IEEE 技术成就奖获得者，Willard 和 Lillian Hackerman 工程讲座主持人，也是马里兰大学巴尔的摩分校（UMBC）计算机科学和电子工程教授。Finin 教授有超过 40 年的人工智能应用于信息系统和语言理解问题的研究经验，目前的研究主要集中在知识图的表示与推理、文本信息的分析与提取、信息系统的安全性与私密性增强等方面。Finin 教授早年曾在 UMBC、Unisys、宾夕法尼亚大学、约翰霍普金斯大学和麻省理工学院人工智能实验室（MIT AI Laboratory）任职。同时，Finin 教授是 400 多篇参考出版物的作者，参与了 DARPA/NSF 的知识共享工作，帮助领导了 KQML 代理通信语言的开发，并且是标准化语义 Web 语言 OWL 的 W3C Web 本体工作组的成员。

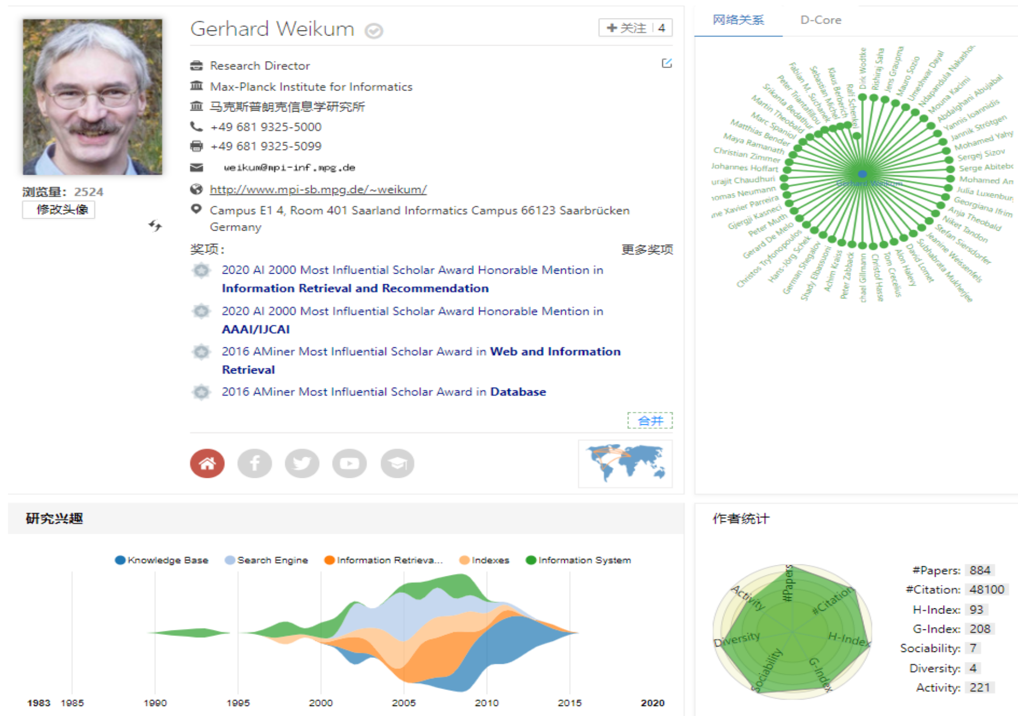
- **Jure Leskovec**



AMiner

Jure Leskovec（h-index：93），是斯坦福大学计算机学院副教授，数据挖掘领域的顶尖学者，在卢布尔雅那大学获得计算机科学学士学位，在卡内基梅隆大学获得计算机科学博士学位。Leskovec 教授的研究重点是挖掘和建模大型社会和信息网络，关注其演变及信息的扩散对其造成的影响。2020 年，Leskovec 教授荣登 AI 全球最具影响力学者榜单。

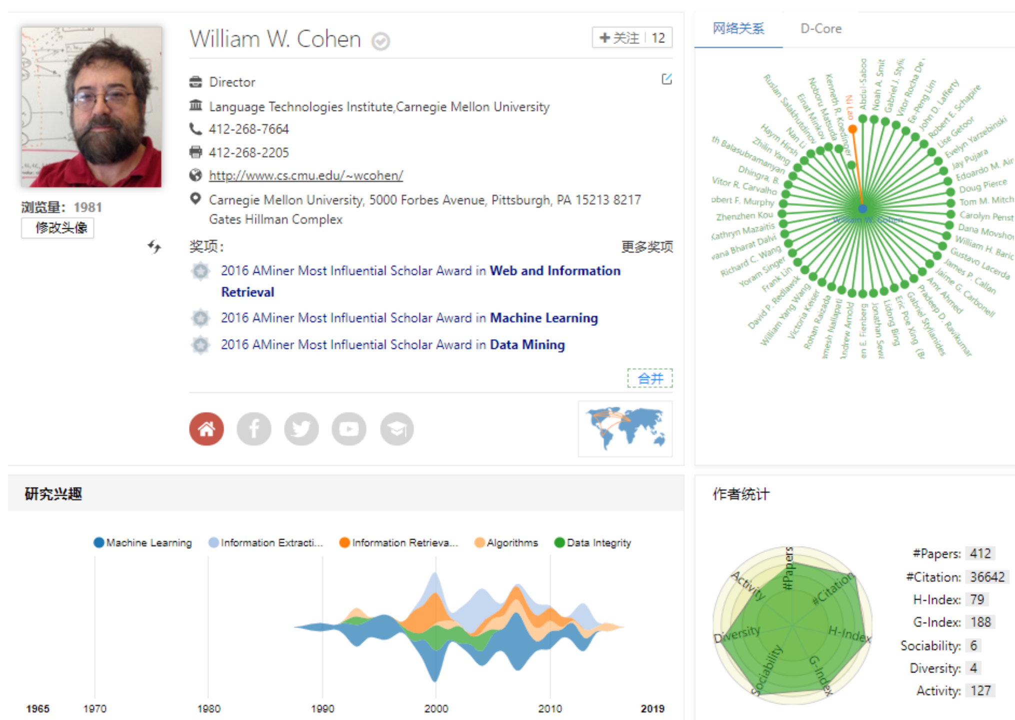
- **Gerhard Weikum**



AMiner

Gerhard Weikum（h-index：93），是德国萨尔州大学计算机科学系的教授，德国 Saarbruecken 马克斯-普朗克信息学研究所（MPII）研究主任，在德国达姆施塔特大学获得博士学位。Weikum 教授早年曾在德国萨尔布吕肯的萨尔大学、瑞士苏黎世联邦理工学院、美国德克萨斯州奥斯汀的 MCC 任职，并在华盛顿雷德蒙德的微软研究院担任访问高级研究员。Weikum 教授的研究方向是分布式信息系统数据库性能优化和自组织数据库与 IR 集成，以及信息提取和知识获取。

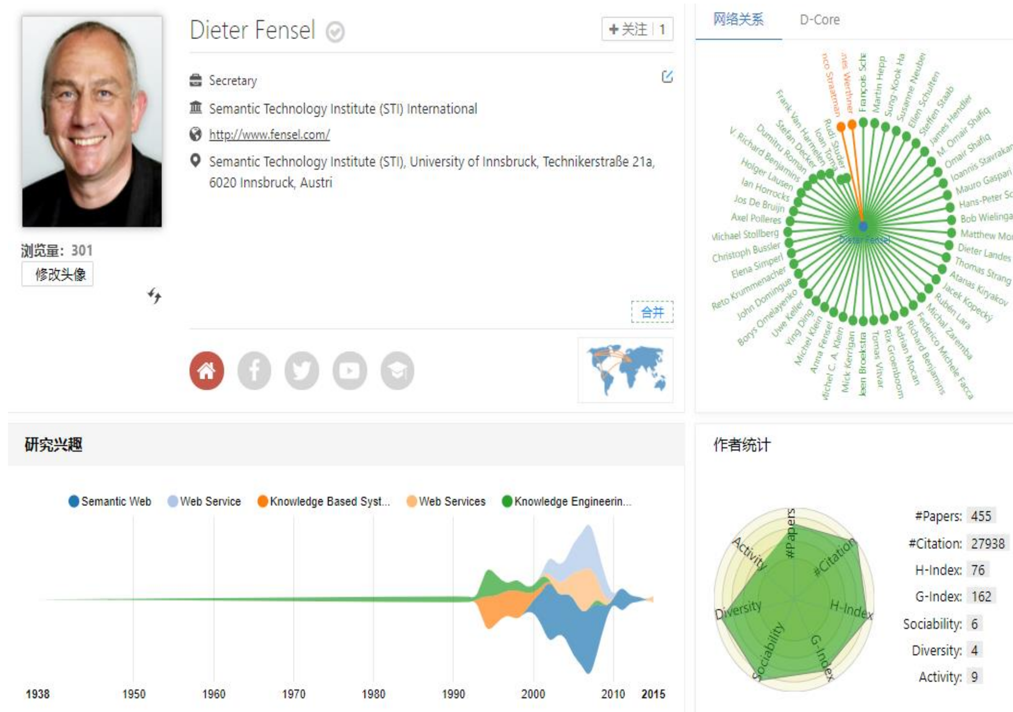
● **William W. Cohen**



William W. Cohen（h-index：79），是 AAAI 研究员，2008 年 SIGMOD“时间测试”奖和 2014 年 SIGIR“时间测试”奖得主，谷歌 AI 的研究与工程总监，在杜克大学获得计算机科学学士学位，在罗格斯大学获得计算机科学博士学位。Cohen 教授曾在卡内基梅隆大学（Carnegie Mellon University）机器学习系工作，并被语言技术研究所（Language Technology Institute）联合聘任为副教授、研究教授和教授。Cohen 教授也曾担任 CMU 机器学习本科辅修课程的主任和 ML 项目理学硕士的联席主任，以及国际机器学习学会的前任会长。Cohen 教授的研究兴趣包括信息集成和机器学习，特别是信息提取、文本分类和大型数据集中学习。共拥有学习、发现、信息检索、数据集成等 7 项专利，发表论文 200 余篇。

AMiner

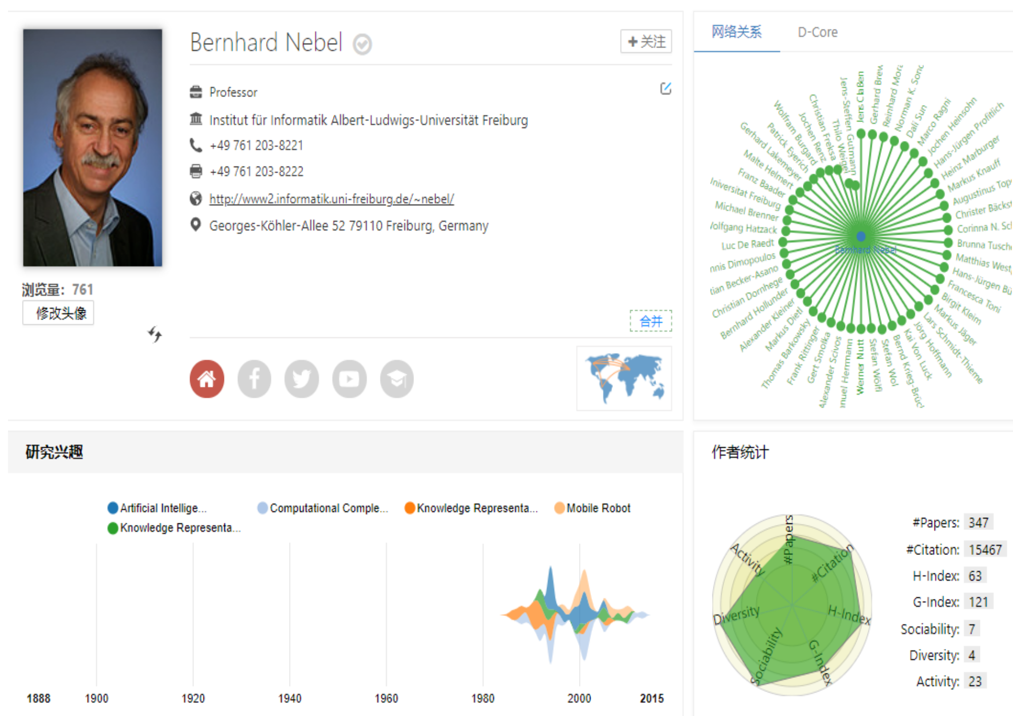
- **Dieter Fensel**



AMiner

Dieter Fensel（h-index：76），是前奥地利因斯布鲁克大学计算机科学教授，在柏林技术大学获得计算机科学硕士学位，在卡尔斯鲁厄大学获得计算机科学学士学位和经济学博士学位。共发表 250 多篇论文，组织 300 多场学术活动，是 14 期科学期刊的联合主编。Fensel 教授曾在卡尔斯鲁厄大学（AIFB）、阿姆斯特丹大学（UvA）和阿姆斯特丹自由大学（VU）任职，在爱尔兰国立大学高威分校（National University of Ireland, Galway）创建数字企业研究所（DERI）并成为科学主任，在奥地利因斯布鲁克大学（University of Innsbruck）创立数字企业研究所（DERI）并成为研究所的创始所长。Fensel 教授还建立了国际语义技术研究所（STI2），这是一个由全球语义网和服务社区中感兴趣的科学、工业和政府团体组成的协作协会。目前，Fensel 教授的研究兴趣是语义在 21 世纪计算机科学中的应用。

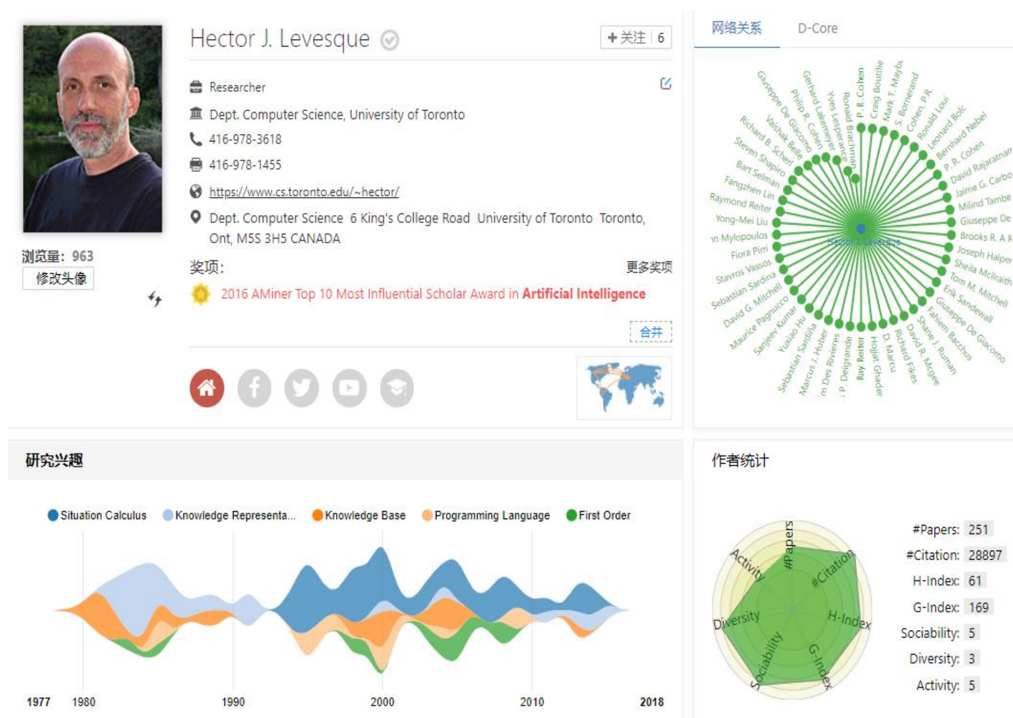
- **Bernhard Nebel**



AMiner

Bernhard Nebel（h-index：63），是 ECCAI 会员，AAAI 会员，德国 Leopoldina 科学院院士，欧洲 paea 研究院成员，阿尔贝托-路德维希-弗莱堡大学（albert - ludwig - universitat Freiburg）教授和人工智能基础研究小组（foundation of Artificial Intelligence）组长，在汉堡大学获得计算机科学博士学位。Bernhard Nebel 教授曾先后在汉堡大学、柏林技术大学、ISI/USC、IBM 德国和德国人工智能研究中心（DFKI）从事不同的人工智能项目，也曾在乌尔姆大学担任副教授（C3），还是合作研究中心 SFB TR/14 复杂系统自动验证与分析（AVACS）的成员和合作研究中心 SFB TR/8 空间认知 Freiburg 站点的协调员。Bernhard Nebel 教授是 10 本书和会议论文集的共同作者和编辑，同时也是 150 多篇科学期刊、书籍和会议论文集的作者和合著者。

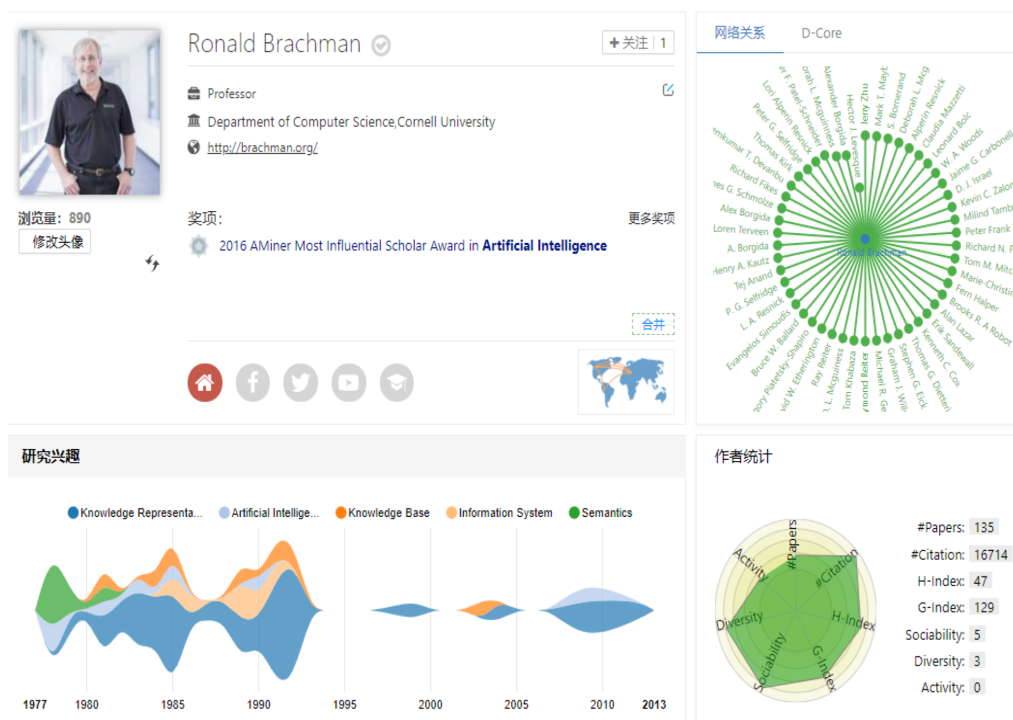
- **Hector J. Levesque**



AMiner

Hector Levesque（h-index：61），是 AAAI 的执行委员会成员，知识表示和推理原理国际会议的联合创始人，在多伦多大学获得理学学士、硕士和博士学位。Levesque 教授的主要研究兴趣为人工智能的知识表示和推理领域，现已发表 60 多篇研究论文和三本书。其中四篇论文分别在 1984 年、1984 年、1992 年和 2006 年获得了美国人工智能协会（AAAI）颁发的最佳论文奖，另外两篇论文在其他会议上获得了类似的奖项。Hector Levesque 教授是五家期刊的编辑委员会成员，其中包括“人工智能”杂志。同时还是第一个获得 IJCAI 颁发的计算机与思想奖的非美国人，以及加拿大自然科学和工程研究委员会 1990-91 年 E.W.R. Steacie 纪念奖学金的获得者。在 1984 年至 1995 年期间 Hector Levesque 教授曾担任加拿大高级研究所研究员，并于 2006 年入选加拿大皇家学会，于 2011 年入选美国科学促进会，于 2012 年获得加拿大人工智能协会终身成就奖。

- **Ronald Brachman**



AMiner

Ronald Brachma（h-index：47），是康奈尔理工学院计算机科学系教授，在普林斯顿大学获得 B.S.E.E.学位，在哈佛大学获得 S.M.和博士学位。Ronald Brachma 教授目前担任 Jacobs Technion-Cornell Institute（以色列理工学院和康奈尔大学的一个创新学术合资企业）的负责人。曾经担任雅虎实验室副主任、首席运营官、负责人和雅虎的首席科学家。雅虎实验室是一个全球性的研究和应用科学组织，拥有一支杰出的科学家和研究工程师团队。

3.2.2 国内代表性学者

- **于旭**



于旭（h-index：72），是 ACM 会员，IEEE 资深会员，IEEE 计算机学会会员，现任香港中文大学系统工程及工程管理学系教授，在日本筑波大学获得计算机科学硕士和博士学位。曾任筑波大学信息科学与电子研究所教员和澳大利亚国立大学计算机科学系讲师，目前主要研究方向为关系数据库关键词搜索、图数据挖掘、图数据查询处理、图数据模式匹配等。于旭教授曾服务于 300 多个组织委员会和国际会议或国际研讨会的项目委员会，担任过 IEEE 知识与数据工程会刊副主编和 VLDB 期刊副主编。目前于旭教授担任亚太网络会议指导委员会主席，并担任“WWW 杂志”、“国际合作信息系统杂志”、“信息处理杂志”和“卫生信息科学与系统杂志”的副主编。

- **唐杰**



AMiner

唐杰（h-index：59），是清华大学计算机系教授和副主任，清华-工程院知识智能联合实验室主任，获杰青，在清华大学计算机系获得博士学位。研究兴趣包括人工智能、数据挖掘、社交网络、机器学习和知识图谱，重点研究方向为设计挖掘社交和知识网络的新算法，共发表学术论文 200 余篇，拥有专利 20 项。唐杰教授曾担任 CIKM'16 的 PC 联合主席，WSDM'15、KDD'18 的副主席，ACM TKDD 的代理主编，IEEE TKDE、IEEE TBD 和 ACM TIST 的编辑。目前担任 AMiner.org 学术社交网络分析与挖掘项目的负责人，该项目已经吸引了来自世界 220 个国家或地区的 1000 多万个独立 IP 访问。唐杰教授曾荣获英国皇家学会牛顿先进奖学金、CCF 青年科学家奖、NSFC 杰出青年学者奖和 KDD'18 服务奖。

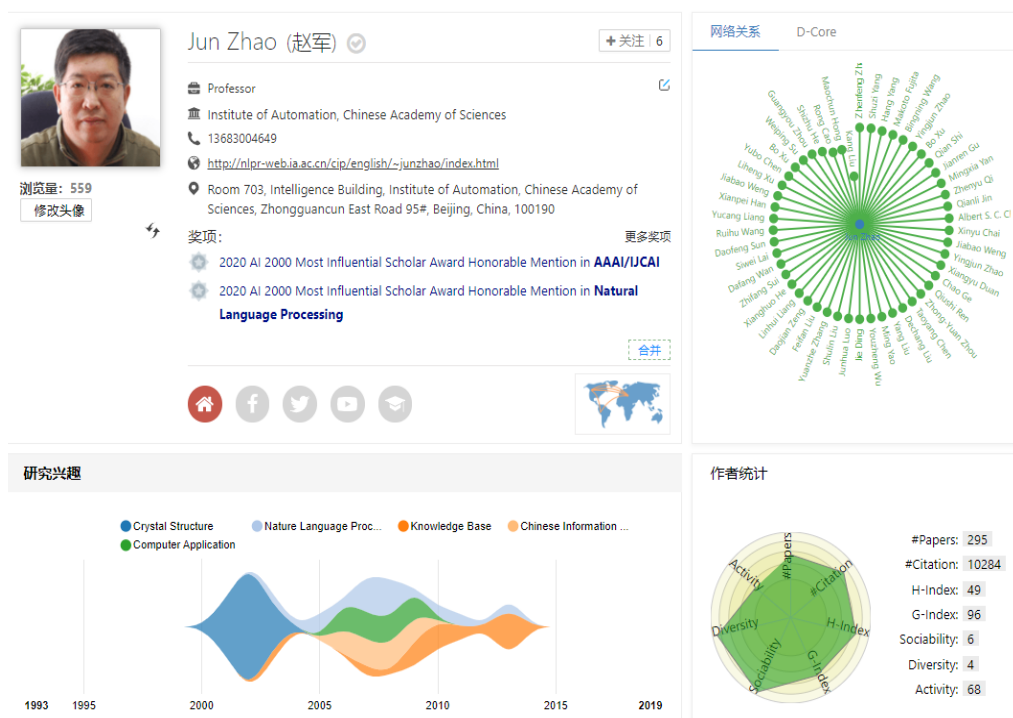
● **文继荣**



AMiner

文继荣（h-index：49），是中国人民大学信息学院教授和博士生导师，国家“千人计划”特聘专家，在中国人民大学信息学院计算机科学与技术专业获得工学学士和硕士学位，在中科院计算所获得博士学位。文继荣教授曾在微软亚洲研究院工作，担任高级研究员和互联网搜索与数据挖掘组主任。工作期间共获得 50 多项美国专利，其中一些成果已经被用于重要的微软产品中（如微软搜索引擎 Bing）。所领导的研究团队开发出了微软学术搜索（http://academic.research.microsoft.com）、人立方（http://renlifang.msra.cn/）、产品搜索等有影响力的互联网应用。文继荣教授曾在国际著名会议和期刊上发表了一百多篇论文，担任过许多国际会议和研讨会的程序委员和主席。目前担任信息检索领域主要期刊 ACM Transactions on Information Systems（TOIS）的副主编（Associate editor）。

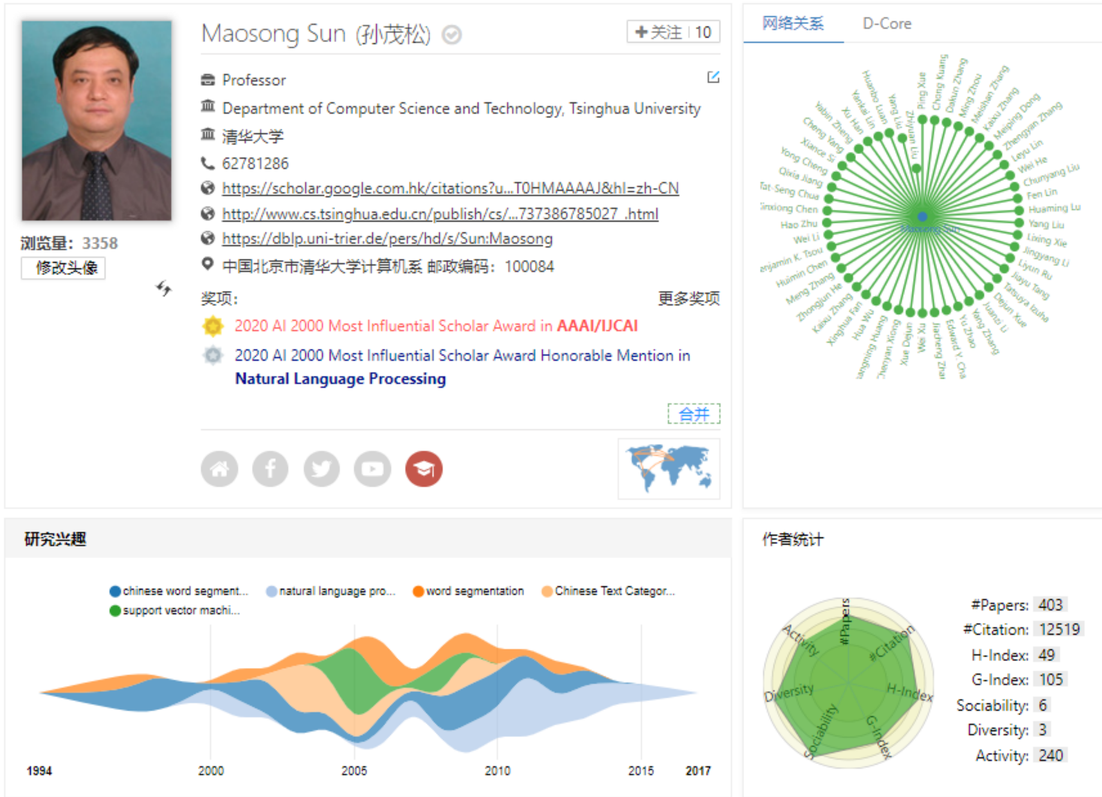
- **赵军**



AMiner

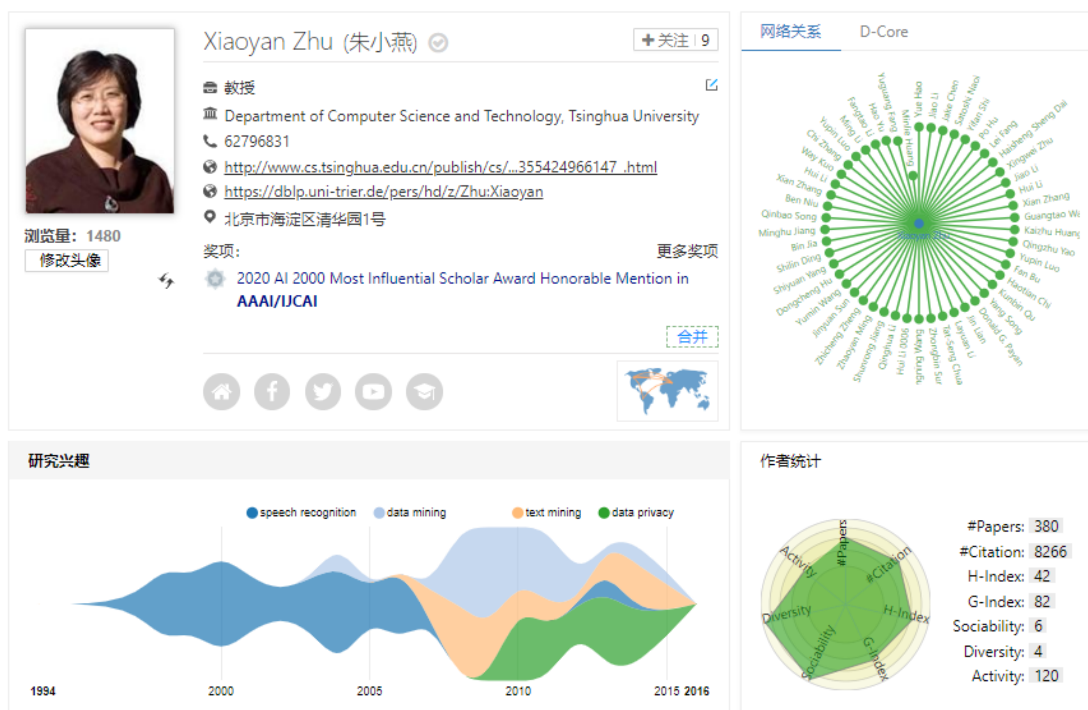
赵军（h-index：49），是中国科学院自动化研究所模式识别国家实验室教授，在清华大学获得计算机科学与技术系博士学位。赵军教授早年曾在香港科技大学担任博士后研究员，也曾经担任 ACL-2016 研讨会的联合主席。目前赵军教授的研究重点是自然语言处理、信息提取和问题回答。赵军教授在国内外著名学术会议和期刊上发表论文 50 余篇，包括 ACL、SIGIR、TKDE、JLMR、IJCAI、EMNLP 等。其中论文“基于卷积深度神经网络的关系分类”获得 COLING-2014 年度最佳论文奖，论文“网络文本中的集体实体链接:一种基于图的方法”在谷歌学术检索中，在 SIGIR 近五年来引用最高的论文中排名第二。

● **孙茂松**



孙茂松（h-index：49），是清华大学计算机科学与技术系教授和博士生导师，清华大学大规模在线教育研究中心主任，教育部在线教育研究中心副主任，清华大学-新加坡国立大学下一代搜索技术联合研究中心共同主任，国家 973 计划项目首席科学家，国家社会科学基金重大项目首席专家。主要研究方向为自然语言理解、中文信息处理、Web 智能、社会计算和计算教育学等。孙茂松教授在国际刊物、国际会议、国内核心刊物上共发表论文 130 余篇，其中在 Computational Linguistics、IEEE Intelligent Systems、ACM TALIP、IJCAI、AAAI、ACL、EMNLP、COLING、VLDB 等国内外一流学术期刊和会议上发表论文数十篇。曾数十次任国际会议或全国性学术会议的大会主席、程序委员会主席或作大会特邀报告。

● 朱小燕



AMiner

朱小燕（h-index：42），是前北京计算机学会副理事长，前清华大学计算机科学与技术系副主任，现智能技术与系统国家重点实验室主任，在北京科技大学获得工学学士学位，在日本神户大学获得工学硕士学位，在日本名古屋工业大学获得工学博士学位。主要研究领域为机器学习，文本信息处理，交互智能和对话系统。曾经担任日本（株）エレクトロダイン研究所研究员，美国加州大学圣芭芭拉分校访问教授，美国康奈尔大学访问教授等，被评为加拿大国际开发研究中心（IDRC）首席科学家。朱小燕教授的脱机手写体汉字与数字识别系统曾获得国家教委科技进步二等奖，第五次全国人口普查光电录入系统获北京市科技进步二等奖。

● 李涓子



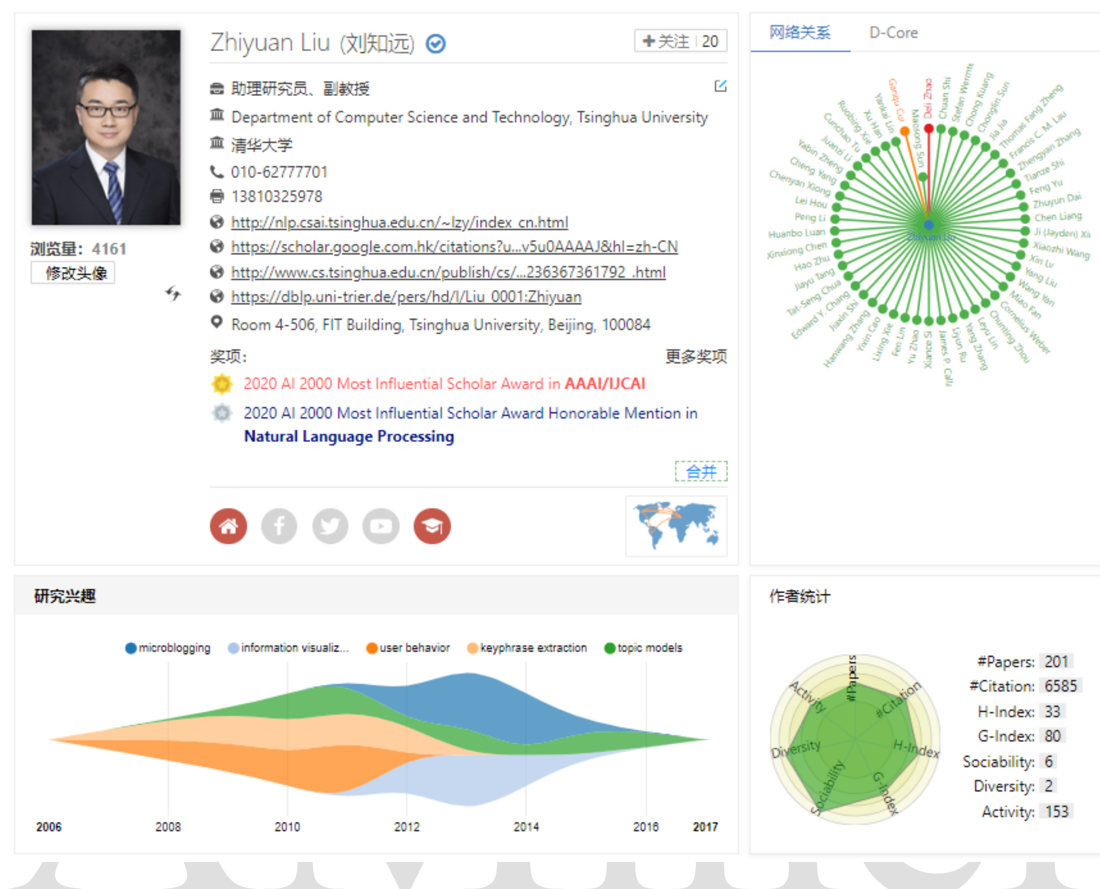
李涓子（h-index：42），是清华大学教授，知识工程组组长，计算机科学与技术学院软件学院副院长，计算机科学与技术学院软件与理论学院副院长，在清华大学获得博士学位。曾任中国计算机联合会中国信息处理技术委员会副主任，亚洲语义网会议（ASWC）指导委员会成员，2006 年 ASWC 地方组织主席。李涓子教授的科研工作特点是融合语义 Web、文本与社会网络挖掘技术，研究基于语义的内容管理关键技术，并应用于包括新闻、研究者社会网络和 Web 服务在内的多个领域。

- **张钹**

AMiner

张钹（h-index：38），是清华大学计算机系教授，中科院院士。曾任美国伊利诺斯大学访问学者，汉堡大学自然科学荣誉博士，校学位委员会副主任，现任微软亚洲研究院技术顾问。张钹教授主要参与人工智能、人工神经网络、机器学习等理论研究，以及这些理论应用于模式识别、知识工程与机器人等技术研究，目前已发表 200 多篇学术论文和 5 篇（或章节）专著（中英文版）。张钹教授发表的专著获得国家教委高等学校出版社颁发的优秀学术专著特等奖，科研成果分别获得 ICL 欧洲人工智能奖、国家自然科学三等奖、国家科技进步三等奖、国家教委科技进步一等奖和二等奖、电子工业部科技进步一等奖以及国防科工委科技进步一等奖等奖励。

- **刘知远**



刘知远（h-index：33），是清华大学计算机科学与技术系副教授，在清华大学获得计算机科学与技术系博士学位。主要研究兴趣包括表征学习、知识图和社会计算。目前刘知远教授已在包括 ACL、IJCAI 和 AAAI 在内的 AI 和 NLP 顶级会议和期刊上发表论文 60 余篇，被谷歌 Scholar 引用 4600 余次。其学术研究成果被评为清华大学优秀博士学位论文、中国人工智能学会优秀博士学位论文。刘知远教授曾荣获清华大学优秀博士后、麻省理工学院 35 岁以下技术创新人才（MIT TR-35 中国）等荣誉，还曾经担任“计算机科学前沿”、“ACL”、“EMNLP”、“COLING”、“IJCNLP”等杂志的青年副主编。

● **肖仰华**



肖仰华（h-index：17），是复旦大学计算机学院教授，在复旦大学获得计算机科学博士学位。肖仰华教授的研究兴趣主要为数据库、数据挖掘、海量数据处理、图形数据库、图形数据挖掘，至今已经在知名国际期刊与会议发表论文 30 多篇。肖仰华教授曾访问美国贝勒医学院人类基因组测序中心，从事海量基因序列数据管理研究，受到微软“铸星计划”支持；访问微软亚洲研究院，从事十亿规模大图数据管理和知识图谱的研究。曾获得中国计算机学会 2010 年度优秀博士论文提名奖、教育部高校科研成果二等奖。

AMiner

3.3 中国学者问题与对策

本报告依托 AMiner 平台，对技术领先的国家、机构和学者，以及认知图谱相关领域的技术发展趋势、技术创新热点等进行了分析，旨在为国家和机构在该领域学科的研究发展和布局提供数据支撑和参考。通过对认知图谱相关领域的论文和学者数据进行统计分析，发现了以下研究问题并给予了相应的对策建议。

（1）我国在认知图谱相关领域的高水平学者主要集中在京津冀、长三角、珠三角、香港和台湾等政治和经济发达地区。我国需要加大经济弱势地区的科研

投入和立项力度，通过科技政策协调各个地区的科研资源配置，改善人才地区分布集中的情况。同时应该加强地区之间的人才合作交流，使领域技术在各地区平衡发展。

（2）美国在认知图谱相关领域的论文发表量和总被引频次、高水平学者数量均位居世界前列。中国虽然在论文发表量和高水平学者数量方面仅次于美国（在全球排名第二），但是与美国相差较大，并且论文总被引频次要低于英国和德国，在全球排名第四。中国相关部门需要重视并采取相关措施，让中国学者在重视论文数量的同时，也要提升论文质量。并且要加快高水平科技人才队伍的建设，一方面针对国际新局势从政策、资金、资源等多个方面加快本土培养，另一方面加大力度引进国外重要机构的杰出学者，改善人才结构，让他们牵头领域技术发展。

（3）针对中国与美国论文合作紧密度较高问题，随着中美贸易战的持续升级，我国应加强与除美国以外的国际交流合作，防止人才限制阻碍中国的发展和产业进步。我国应该提供更多本土学者与其他国家的合作机会，并增加合作政策的宣传力度，吸引国外杰出人才和顶尖机构参与合作。同时，对于合作情况可观的国家可达成长期合作，扩大合作优势。

AMiner

AMiner

4　应用篇



4 应用篇

机器认知智能的发展过程本质上是人类脑力不断解放的过程，是人工智能的最高阶段。认知图谱是实现机器认知智能的底层支撑手段，它不是束之高阁、高高在上的前沿技术，而是一类能够实实在在落地的、有着广泛且多样的应用需求的、能够产生巨大社会经济价值的技术，应用场景主要体现在电商平台、智慧城市、司法行业、金融行业、安防行业、精准分析、智慧搜索、智能推荐、智能解释、更自然的人机交互和深层关系推理等各个方面[182]。

4.1 电商平台

电商平台最大的挑战是从日益增长的海量商品（数十亿）中挑选出的一个小的子集（几十或上百）展示给用户，以满足用户的个性化的购物需求。建设大规模电商认知图谱可以有效解决重复推荐、缺少新意等问题。

通过构建电商认知知识图谱，将用户-场景-货物进行有效的关联，可以挖掘出更多的用户标签，精准感知用户场景，从而使电商搜索和推荐更加准确，有效提高货品转化率。以更好地认知用户需求为目标的电商认知图谱，将助力搜索推荐等从基于行为的方式迈向基于行为与语义融合的认知智能时代，将是平台生态稳定和日益进步的重要基础。

电商相关的文本语料稀缺，品类词在语料中的共现非常稀疏，给抽取带来了极大的难度。针对这些难点，阿里巴巴正在着手设计一套人工+算法不断迭代优化的 active learning 流程，希望为后续的 concept 理解和推理应用提供可靠的支持。阿里巴巴集团搜索推荐事业部认知图谱团队，旨在打造全球最大的中文电商知识图谱，支持包括淘宝、天猫优酷乃至海外电商在内整个阿里集团的推荐与搜索业务，每天服务上亿用户。电商"认知"图谱，从电商场景下的用户需求出发，不局限于传统的商品图谱，而是一个连接商品，用户，购物需求，以及各类开放领域知识、常识的大规模语义网络。生长于全球最大的中文电商平台，认知图谱赋能搜索、推荐等核心业务，是推动集团新零售战略的强大知识引擎。

下面以阿里巴巴电商平台作为分析案例，从认知推荐、基础数据层、推理引擎层、用户交互的文本和视觉智能等方面，详细介绍认知图谱如何赋能电商平台

AMiner

的搜索、推荐等核心业务。

4.1.1 认知推荐

2019 年，全球零售电子商务销售额达 3.53 万亿美元，电子零售收入预计到 2022 年将增长至 6.54 万亿美元。如此快速的增长为全球电子商务行业带来广阔的前景，这标志着一个强劲的市场和广阔的客户需求。除了流量的巨大增长外，各种即将到来的模式也在迅速增长，包括短视频，直播，达人推荐等。随着新出现的各种模式以及消费者对于推荐系统的更高要求，必须要更加系统的解决对于消费者需求的认知推理工作。基于此，我们在全球月活用户和流量最大的电子商务平台手机淘宝的推荐系统上，不断打磨和落地第二代 AI 系统认知智能计算平台。该认知智能计算平台包括三个主要模块，基础数据层，推理引擎层和用户交互层。其中基础数据层，我们主要专注于：（1）跨领域知识图谱的构建；（2）拉通跨域跨场景各类行为数据，例如浏览、点击、收藏、加购、转发等，全方位建模经济体内行为，差异化理解多模态行为数据背后的含义；（3）在意图感知的实时性与认知深度间做平衡，层次化强化消费者理解。推理引擎层，我们依托于（1）多模态预训练和理解对商品理解和进行全域召回，缓解马太效应，加强手淘生态建设；（2）超大规模图神经网络系统进行消费者意图推理。用户交互层，我们通过用户交互的视觉智能和文本智能，通过短视频改变和引导购后消费者心智，理解消费者意图，助力消费者决策，同时填补填补目前学界和业界空白的基于用户交互的弱监督内容理解方向。

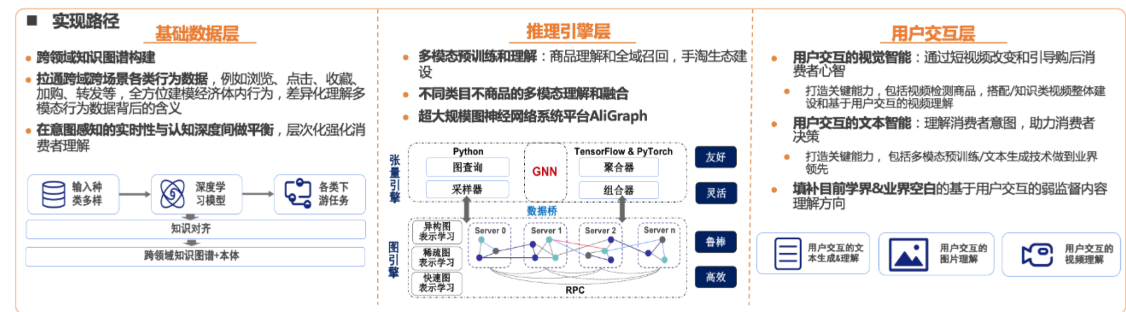


图 52 阿里巴巴认知智能计算平台

4.1.2 基础数据层

阿里巴巴生态里积累了海量的跨领域多模态（商品、短视频、直播等）数据，

这些宝贵的商品数据来自于淘宝、天猫、1688、AliExpress 等多个市场，同时品牌商、行业运营、治理运营、消费者、国家机构、物流商等多种角色参与其中，贡献着校正着这样一个庞大的多模态数据库。但无论从知识产权保护角度，还是提升消费者购物体验，实现跨领域多模态数据的标准化（跨领域多模态规范的统一和跨领域多模态信息的确定性），以及与内外部数据之间的深度互联，意义都非常重大。阿里跨领域知识图谱承载着跨领域多模态标准化这一基础性、根源性的工作。通过知识表示来规范对多模态数据的描述，这样我们才能知道比如哪些商品是同样一件产品，才能确切地知道一个品牌是否被授权，品牌下的产品卖到了哪些市场。

此外，阿里知识图谱综合利用前沿的内容理解、语义推理和深度学习等技术，打造全网商品智能服务体系，服务阿里生态中的各个角色。跨领域知识图谱广泛地应用于搜索、前端导购、平台治理、智能问答、品牌商运营等核心、创新业务。能够帮助品牌商透视全局数据，帮助平台治理运营发现假货，帮助行业基于确定的信息选品，做人货场匹配提高消费者购物体验等等。为新零售、国际化提供可靠的智能引擎。

AMiner



图 53 阿里巴巴数据生态

4.1.3 推理引擎层

推荐系统的本是系统用户与系统内容之间的个性化匹配。我们也可以认为每一次匹配是对用户当下状态的一种“推理”。这里的认知主要体现在用户，即人类的参与。推荐系统需要通过用户在系统中留下的“足迹”，来对用户当下的认知状态进行猜测，从而提供合适的内容。

用户的“足迹”是由人类的认知所驱动的。我们可以想象每一个用户背后都有一个符合认知理论的大脑，驱动着 TA 在系统中的行为，例如接受信号以及给与反馈。推荐系统能否通过用户的反馈，来试图揣测甚至还原用户对于系统的认知状态，是非常有意思的问题。

（1） 用户表征

我们首先希望有一个算法，能够通过这类反馈来表征这个用户。假设用户在系统中有 n 类行为，每一类行为我们都可以用 < 动作类型 a，目标 o，时间 t > 的三元组来表示。动作类型可以是点击/收藏/加购、领取/使用等等。目标可以是商品，优惠券，搜索词等等。我们接下来如何用这一系列的三元组来表示用户呢？

这显然可以被认为是一个序列建模的问题。用户序列的建模通常会用 RNN（LSTM/GRU）或者 CNN + Pooling 的方式进行。然而 RNN 难以并行，训练和预测时间较长，且 Internal Memory 无法记住特定的行为记录。例如 NTM 中就给出实验说明 LSTM 甚至无法完成简单的复制任务。CNN 则需要多层才能使距离较远的行为互相影响。

试想不同用户的相同行为，他们原始的表征应该是一样的，因为行为三元组的输入完全相同。然而，在不同用户的记忆里他们扮演的角色却可能千差万别。这是因为每个行为在人大脑中的记忆并不是孤立的，而是受历史影响，未来也会持续更新的。相比于 RNN/CNN 而言，Self-Attention 完全符合上面的直觉，且有着令人惊讶的记忆法则：经过一次 Self-Attention 操作之后，重复的事件表征会被模糊，出格的事件反而会被保留的比较清晰。这一点有趣的性质，促使我们希望使用 Self-Attention 来对用户行为进行建模——因为看上去是比较符合认知，

AMiner

或者至少有往后继续探索认知的潜力的。而接下来的网络结构则成为了今后用户表征的一个常见做法。

- **ATRank 的网络结构**

整个用户表征的框架包括原始特征层，语义映射层，Self-Attention 层和目标网络。语义映射层能让不同的行为可以在不同的语义空间下进行比较和相互作用。Self-Attention 层让单个的行为本身变成考虑到其他行为影响的记录。目标网络则通过 Vanilla Attention 可以准确的找到相关的用户行为进行预测任务。使用 Time Encoding + Self-Attention 的思路，我们的实验表明其的确可以替代 CNN/RNN 来描述序列信息，能使模型的训练和预测速度更快。

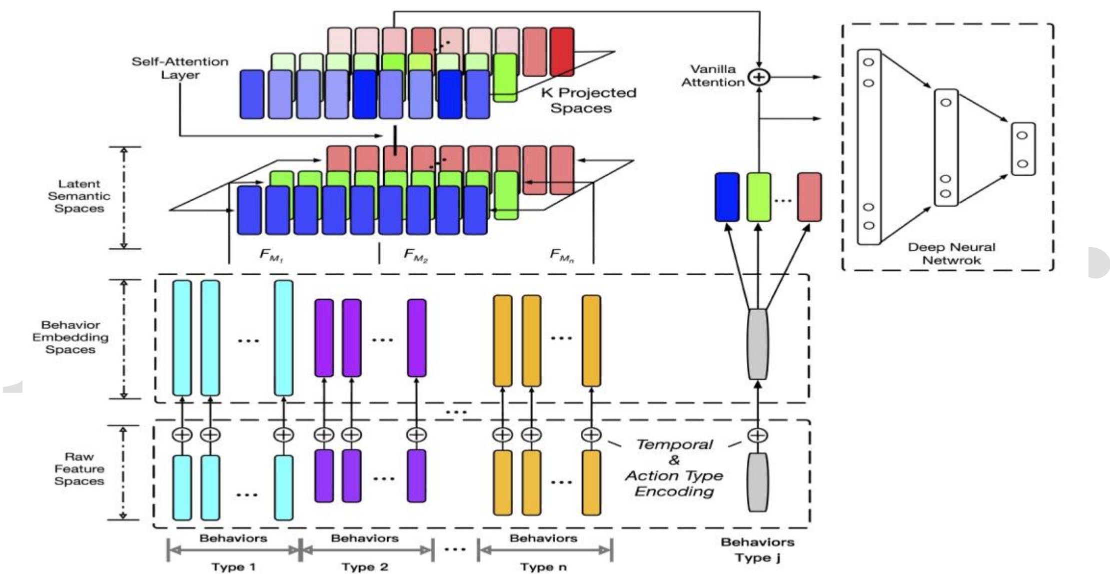


图 54 ATRank 的网络结构

- **行为分组**

某个用户的行为序列可以用一个三元组来描述<动作类型 a，目标 o，时间 t>。我们先将用户不同的行为按照目标实体进行分组，如图中最下方不同颜色 group。

对三元组的行为编码过程如下：每个实体都有自己不同的属性，包括实值特征和离散 id 类特征。动作类型是一个 id 类特征，我们也将时间离散化成一个 id 类特征。用户可自行对目标实体进行编码，再与动作和时间的 embedding 三部分相加得到下一层的向量组。

即，某行为的编码 = 自定义目标编码 +lookup（离散化时间） +lookup（动作类型）。

由于实体的信息量不同，因此每一组行为编码的向量长度不一，其实也代表行为所含的信息量有所不同。另外，不同行为之间可能会共享一些参数，例如店铺 id，类目 id 这类特征的 lookup table，这样做能减少一定的稀疏性，同时降低参数总量。

分组的原因除了说明起来比较方便，更主要还是实现原因，因为我们认为变长、异构的处理很难以在不分组的情况下进行简单、高效的实现。而在后面可以看到我们的方法实际上并不强制依赖于行为按时间排序。

- **语义空间映射**

这一层通过将异构行为线性映射到多个语义空间，来实现异构行为之间的同语义交流。例如框架图中想表达的空间是红绿蓝（RGB）构成的原子语义空间，下面的复合色彩（不同类型的用户行为）会投影到各个原子语义空间。在相同语义空间下，这些异构行为的相同语义成分才有了可比性。

尽管从实现的角度上来说，这一层就是所有行为编码向一个统一的空间进行映射，映射方法线性非线性都可以，但实际上，对于后面的网络层来说，我们可以看作是将一个大的空间划分为多语义空间，并在每个子空间里进行 self-attention 操作。因此从解释上来说，我们简单的把这个映射直接描述成对多个子语义空间进行投影。

- **Self-Attention 层**

Self-Attention 层的目的实际上是想将用户的每一个行为从一个客观的表征，做成一个用户记忆中的表征。客观的表征是指，比如 A,B 做了同样一件事，这个行为本身的表征可能是相同的。但这个行为在 A，B 的记忆中，可能强度、清晰度是完全不一样的，这是因为 A,B 的其他行为不同。实际上，观察 Softmax 函数可知，某种相似行为做的越多，他们的表征就越会被平均。而带来不一样体验的行为则会更容易保留自己的信息。因此 Self-Attention 实际上模拟了一个行为被其他行为影响后的表征。

AMiner

另外，Self-Attention 可以有多层。可以看到，一层 Self-Attention 对应着一阶的行为影响。多层则会考虑多阶的行为影响。这个网络结构借鉴的是 Google 的 Transformer。

- **目标网络**

目标网络会随着下游任务的不同而定制。论文里用的是行为预测推荐的任务，采用的是 point-wise 的方式进行训练和预测。

灰色的 bar 代表任意种类的行为。我们将该行为也通过 embedding、projection 的转换，然后和用户表征产出的行为向量做 vanilla attention。最后 Attention 向量和目标向量将被送入一个 Ranking Network。其他场景强相关的特征可以放在这里。

- **实验与 case 分析**

在行为预测或推荐任务中，我们通过实验发现 self-attention + time encoding 也能较好的替代 cnn + pooling 或 lstm 的编码方式。训练时间上能较 cnn/lstm 快 4 倍。效果上也能比其他方法略好一些。

AMiner

（2） 解离化用户—商品表征学习

传统表征学习算法得到的用户商品表征本身，并不具备可解释性，而往往只能提供用户-商品之间的 attention 分作为商品粒度的用户兴趣。能否仅通过用户行为，学习到本身就具备一定可解释性的解离化的用户商品表征，并试图利用这样的商品表征完成单语义可控的推荐任务？

什么是解离化表征？

学习解离化表征的原因和人类认知的 Binding Problem（BP）[3]有关。如上图所示，当我们看到一副上面四张图片时，我们可以识别几种较为独立的语义元素，即颜色、形状和位置，这个过程对应 Segregation Problem（BP1）。如果我们学到的表征向量中，不同的维度能够分别代表不同的语义，我们称这样的表征是 Disentangled 的，反之如果某一维对应多种语义，则称该表征是 Entangled 的；而我们通过这些识别独立的语义元素进而可以组合生成下面四张新的图片，这个过

程对应 Combination Problem（BP2）。这里假设我们有一个生成模型，能通过 BP1 学到图片的 Disentangled 的表征，我们通过调整表征中单个维度的值即可达到只改变某个语义而不影响其他语义的生成，例如把绿色的方块变成绿色的心形，颜色位置不变。除了和人认知过程比较切合，还可以提供可控制的生成之外，学习解离化的向量表征的鲁棒性更好，对测试数据的分布相比而言敏感度较低，同时也有较好的解释性。

用户-商品解离化表征：

在电商系统中，拥有大量的用户、商品标签以及庞大的用户行为数据，如果可以通过这些数据学习解离化的用户和商品表征，将对推荐可解释、认知推理、营销等工作产生非常积极的影响。

而我们此次的工作，是用户-商品解离化表征工作的一个开端，因此我们仅使用到用户行为数据，并不涉及任何商品属性以及用户特征。我们探索，从用户行为当中，能否得到一些认知相关的决策因素并以可解离的方式对商品和用户进行表示。

AMiner

具体来说，我们关注这样的两个和认知相关的子任务：

a）商品在人的认知空间中，他们是如何表征的，这样的表征是否具有可解释性，例如是否能找到的对应的某一维就能够代表一个独立的“语义”。类似的，人在这个空间下的表征，是否也具有这样的语义。联系解离化表征（Disentangled Representation Learning）在图像数据上的发展，我们探索其是否能从离散数据，特别是用户行为数据上学习到类似的结果。

b）基于这样的表征，我们能否提出新型应用，并至少给出一种原型方案。

- **MacridVAE:**

针对电商平台上用户行为的特点，我们的模型采用了层次化的设计：它在推理一个用户的表征时将依次进行宏观解离化（Macro Disentanglement）和微观解离化（Micro Disentanglement）。

- **宏观解离化：**

宏观解离化的启发一方面来自于用户在综合类电商场景下的兴趣确实较为分散，另一方面也来自于人类的传统认知难题 Binding Problem。首先，用户兴趣通常是非常广泛的，一个用户的点击记录往往会涉及到多个独立的消费意图（比如点击不同大类下的商品），也有人把这个叫多峰的兴趣。而用户在执行不同意图时的偏好往往也是独立，比如喜欢深色的衣服并不意味着用户也喜欢深色的电器。哪怕是价格偏好也常存在不可迁移的情况，比如买高档口红、和买便宜好用的笔记本电脑这两者并不互斥。另外，宏观解离化也是微观解离化的必要前提（见下）。因此，我们的将用户的表征拆分成 k 个 d 维分量。

- **微观解离化：**

我们希望能把用户在执行某个意图时的偏好进一步地分解到更细的粒度。比如，设第 k 个意图对应服饰，我们希望用户在这个意图下的偏好向量 的各个维度能够对应不同的商品属性，比如某一维和颜色相关，另一维和尺寸相关，等等。这里我们看到，宏观解离是微观解离的前提：不同大类的商品属性集合是很不同的，用户表征向量的某一个维度如果已经被用于刻画用户对手机电量的偏好了，那么这一维对服饰等商品就是没有任何意义的——在预测用户是否会点击某个服饰时、在通过用户行为学习某个服饰的表征时，我们都应当忽略这些只和手机相关的维度。

- **从用户行为学习解离化表征：**

我们给出模型实现架构总览如下图：

注意前半部分的 Macro Level 的用户偏好向量挖掘可以用任何带“聚类”功能的前向网络实现，例如 neural EM，capsule network，mixture of expert 等。经历大量的尝试和调整，形成了最终论文中的编码网络实现——由一组全局的中心向量来实现聚类的方式。

AMiner

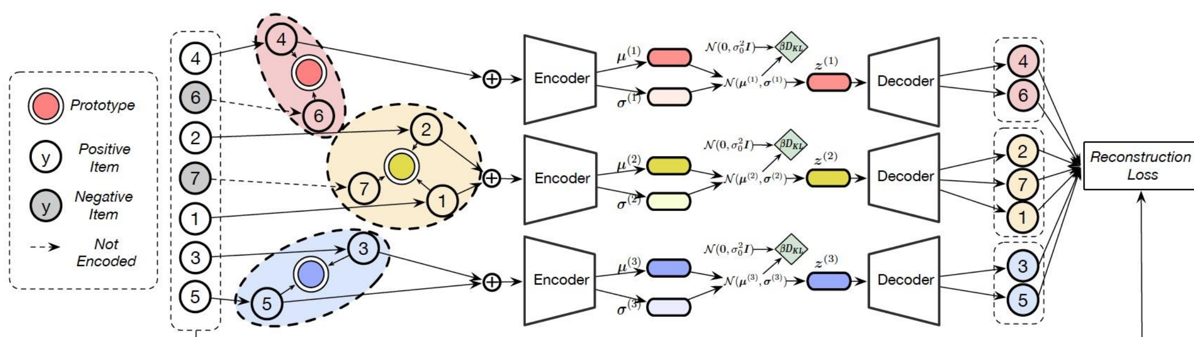


图 55 从用户行为学习解离化表征模型实现架构总览图

为了优化这个深度概率模型，我们采纳了 VAE 的框架。这里我们为了实现微观解离化，借鉴了 beta-VAE[1] 的方法，对 KL 惩罚项进行了加强（令 beta 远大于 1）。这种做法将迫使表征的各个维度去捕捉比较独立的信息，比如各种各样的不太相关的商品属性（颜色和大小就不太相关）——当然，某些商品属性之间可能存在较强的相关，这就有待未来工作去解决了。

- **用户可控制的交互式推荐**

AMiner

解离化表征在带来一定的可解释性的同时，也带来了一定的可控制性。这种可控制性有望给推荐系统引入一种全新的用户体验。比如说，既然表征的各个维度关联的是不同的商品属性，那么我们完全可以把用户的表征向量提供给用户，允许用户自行固定绝大部分维度（比如对应的是衣服的风格、价格、尺寸等）、然后单独调整某个维度的取值（比如颜色对应的维度），系统再根据这个反馈调整推荐结果。这将帮助用户更加精准地表达自己想要的、并检索得到自己想要的。

我们可以从某个商品或用户的表征出发，在固定住其它维度后，我们逐步地改变表征第 j 维的取值。然后我们利用论文中提供的一种类似 beam search 的小技巧，检索出了表征第 j 维大不一样、但其它维度都很一致的一批商品。

我们在淘宝数据集上进行了单语义控制的 case study。以下是我们在调控某个维度后检索得到的两批商品，可以看出这个被调控的维度和背包的颜色这一属性比较相关，并有较为明显的渐变性质：



图 56 背包颜色维度示例图

以下是我们在另一个维度上检索得到的两批商品，可以看出这个维度和背包的大小这一属性比较相关，并有较为明显的渐变性质：



图 57 背包大小维度示例图

然而，目前并不是所有的维度都有人类可以理解的语义。而且，正如文献[2]所指出的，在无监督的情况下，训练出可解释的模型仍然是需要运气的——在加了 beta-VAE 的约束后，获得可解释模型的概率相比普通 VAE 大大提高，但仍然避免不了“反复训练多个模型，然后挑出最好的模型”这一陷阱。因此，研究者们应多多关注（弱/半）监督方法，引入标签信息。

AMiner

4.1.4 用户交互的文本和视觉智能

（1） 基于用户交互的自然语言生成

自然语言处理中很多任务依然面临巨大的挑战，而其中一大挑战便是如何实现基于用户交互的自然语言生成。用户交互信息繁多复杂，文本生成如何有效利用这类信息生产更加符合用户兴趣的文案是一个具有重大技术价值和业务价值的问题。比如在商品推荐场景中，自动生成的推荐理由能够为用户提供更多商品相关的信息，生动的描述更容易提升用户的体验，激发用户的兴趣和提供更好的“逛街”乐趣。推荐场景的文本生成更加复杂，需要充分考虑文本生成系统和用户之间的交互信息，从而实现根据用户交互信息提升文本生成的用户体验。因此推荐场景的文本生成不仅要求系统生成自然流畅的文案，同时还需要生成让用户感兴趣和满意的文案。

而在阿里巴巴淘宝的推荐场景中，云主题便是一个需要自动化实现推荐理由生成的应用。云主题是由一系列商品所形成的主题场景，其中往往包含特定生活场景所需的商品，如郊游、洗漱、电竞等等。算法团队利用现有推荐系统中的召回技术实现基于场景的商品聚合，然而面临如此繁多的主题，利用人工生成标题和推荐理由将消耗非常大量的人力物力。面临这个需求，阿里巴巴算法团队则利用自然语言生成技术做到自动化产出标题和推荐理由。这项工作主要有以下几个难点：1. 生成的文案需要符合语法，表达流畅通顺；2. 推荐理由的内容需要与商品属性相关，同时要符合商品相关的知识，比如食品的推荐理由不能使用其他领域的表述，比如"便于组装"这类常用于描述玩具的表达；3. 推荐场景下的文本生成要尽可能做到个性化，对于同一个商品类目，不同的用户群体能够获得不同的推荐理由。

过去文本生成大多基于传统的基于 RNN 的 Seq2Seq+Attention 的框架实现。这个模型最早用于机器翻译，后来逐步扩展到各大文本生成类的任务。而当前机器翻译领域已经发生了一些改变，学术界和工业界普遍青睐于速度更快、效果更好的 Transformer 架构。如今 Transformer 架构已被引入到推荐理由生成的场景中，在 baseline 的层面上获得比传统的 RNN-based Seq2Seq 更好的表现。

而为了让生成质量和用户体验均获得提升，算法团队引入了外部知识和个性化信息。为了实现推荐理由的知识引入和个性化，算法引入了特征属性和知识。特征属性指的是商品的某些方面，如外观、质量等。针对每一个商品，当前的输入除了商品标题以外，还有与它对应的特征属性。此外，算法使用了外部知识库 CN-DBpedia，使得输入的商品和其相关的知识联系起来。

针对以上需求，算法团队设计了相应的 KOBE 模型（Knowledge-Based Personalized text generation system）。该模型建立在基于自注意力机制的 transformer 模型，下图展示了 KOBE 模型的基本框架：

编码器由多层的自注意力层叠而成（一般为 6 层，如果要实现大模型的话可以增加层数）。编码器接受文本序列的输入，通过词嵌入层将其转变为词嵌入序列。而由于模型不同于 RNN，无法直接捕捉到输入的序列关系，算法给词嵌入序列加上位置编码以后再送入编码器的自注意力层。每个自注意力层主要包括自

AMiner

注意力机制和全连接层，内部使用了残差连接和 layer normalization。在自注意力模块后，全连接层（FFN）将之前的输出进行变换。而解码器的实现近似于编码器，同样采用了多层的自注意力层。不同的是解码器增加了 context attention，将编码器和解码器联系起来。

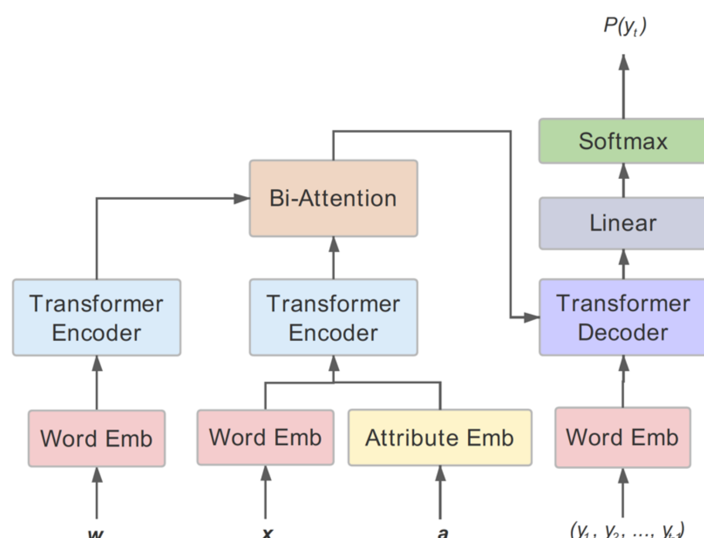


图 58 KOBE 模型的基本框架

AMiner

在 Attribute Fusion 模块中，算法引入个性化推荐相关特征属性，主要集中在两个方面。第一个是 aspect，指的是产品的特定方面，比如外观、质量等。第二个是 user category，即该产品匹配的用户群体，比如机械键盘对应的用户群体是科技爱好者等，算法利用 Attribute Fusion 模块将它们实现有机融合。另一个文本生成的常见问题就是缺乏信息量，算法设计了 knowledge incorporation 这一模块。模块中的知识编码器同样是一个基于自注意力机制的编码器，具体模型构造和上文所述编码器一致。为了实现商品标题表示和相关知识表示的有效结合，算法实现双向注意力机制，分为“标题到知识的注意力”和“知识到标题的注意力”。两者的功能各有不同。标题到知识的注意力能够获取和标题相关的知识，而知识到标题的注意力则获取和知识相关的标题内容。

阿里巴巴算法团队在春节云主题的项目上线了 KOBE 模型，为每个类目提供相应的个性化推荐理由。可以看到生成的推荐理由表达流畅，能够根据类目和产品特点选择符合产品相关知识的描述，并且能够携带一些具有趣味性的表达。



图 59 商品个性化推荐示例图

（2） 多模态预训练实践

随着全球零售电商销售规模的快速增长，以及电商系统与用户交互形式日益多元化，电商系统理解和检索图片、视频等多模态内容的需求日益增强。面对深度多模态表征学习的需求，阿里巴巴算法团队跟进多模态预训练的相关工作，设计了一套新的多模态预训练方法，学习更加高质量的多模态表征，并将其应用到电商推荐场景中，实现搜索 query 召回相关商品图片以及多模态商品召回等相关任务。尤其在商品召回场景，如何充分结合多模态特征和用户交互信息从而实现效果更优的推荐，则是非常值得探索的课题。

算法自研设计了一套多模态表征学习的框架 InterBERT，模型框架如下：

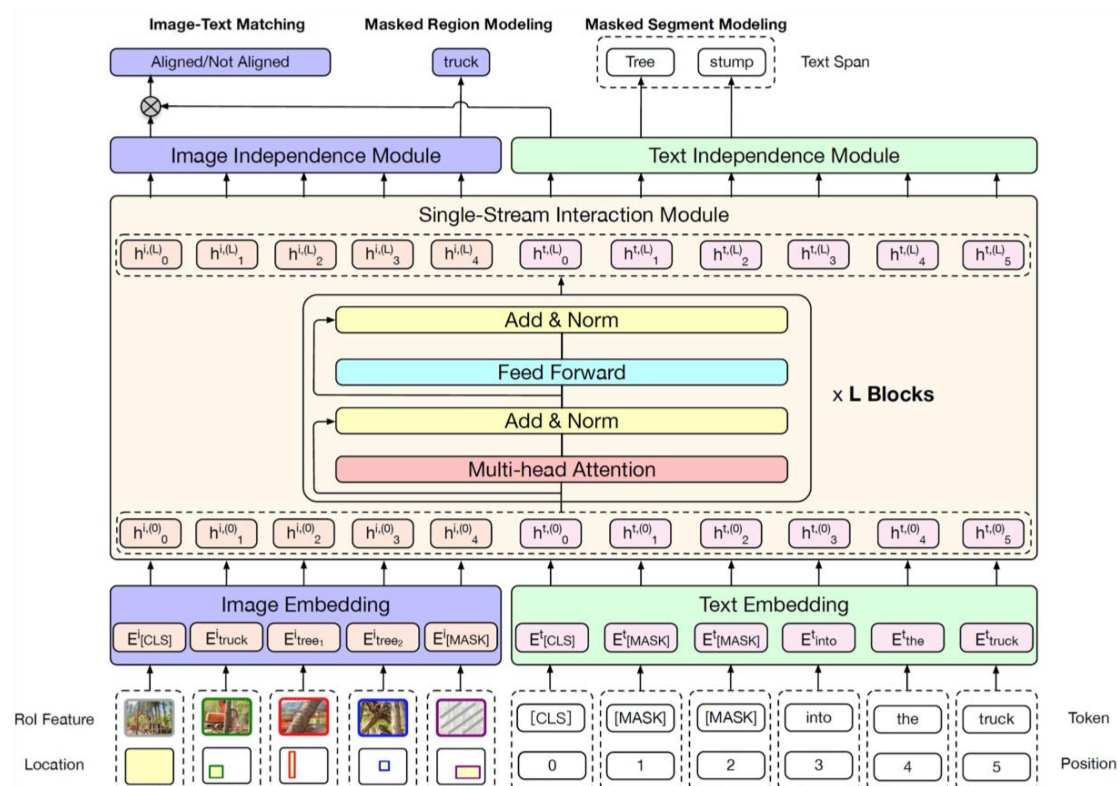


图 60 多模态表征学习框架

AMiner

输入 query 序列和商品图片物体序列，经过 InterBERT 模型的 interaction module 和 independence module 编码，最后预测 query 与商品图是否匹配，文本侧用 BERT 初始化。

多模态预训练主要从 2019 年开始有了比较大的进展，几项比较有代表性的工作都是参考了自然语言处理的预训练，将 Transformer 架构迁移到多模态预训练中，并做了相应的修改。多模态预训练的数据主要是图文对，如图片和它对应的描述（caption）。在这里文本可以看成是一个字/词/subword 序列，和 Transformer 模型的输入一致。但是对于图像，要想和文本一起送入 transformer 模型，最合适的方法就是把图像也变成一个序列。于是前人便非常巧妙地借助 object detector，首先将图像中的物体检测出来，然后再把物体的表征和物体对应的位置以 embedding 的形式送入模型，实现跨模态学习。而在训练中，文本侧依然采用经典的 masked language modeling，而图像侧则借助 detector 得到的物体的类别，采用 masked object modeling，让模型在观察上下文预测被 mask 的物体所属类别。而 BERT 的 Next-sentence prediction 任务则转变成了 image-text pairing 的任务，只需人为地构造图文不匹配的负例便可实现该任务的训练。

而模型架构方面，主要包含两种流派，一种是 single-stream 的模型，如 Visual BERT，UNITER 和 Unicoder-VL，这类模型将文本图像特征拼接在一起一并送入 Transformer 模型；另一种则是 two-stream 模型，如 VilBERT，VL-BERT 和 LXMERT 等，这类模型将文本和图像分别编码，在中间过程利用 co-attention 的架构实现跨模态交互。事实上，两种模型各有优劣势，single-stream 模型直接将两种模态融合在一个模型中，交互更加充分，但每个模态的表示缺少自身独立性；two-stream 模型每个模态具有良好的独立性，但是交互过程采用 co-attention 每次只能看到另一个模态的信息而不能采集自身模态的上下文信息，一个 ad-hoc 的手段是做完 co-attention 以后做 self-attention，这样除了效果不直接以外参数量也大幅增加。

参考了两类模型的优缺点，算法团队设计了一套结合两者优势的模型结构 InterBERT，其中包括 single-stream interaction module 用于更好的跨模态交互以及 two-stream independence module 来保持模态的独立性。具体实现如模型图所示，输入的表征首先经过 single-stream 模型实现交互，这种交互每次 self-attention 能够充分观察两个模态的上下文信息，从而实现跨模态的理解，而在输出前使用 two-stream 的架构保证输出的每个模态的表征具有自身独立性，模型在交互后对各个模态的信息进行进一步的单独编码。而在训练方法方面，算法团队也对训练任务做了一些改进，让模型建模图像物体和文本序列的能力更强，主要是提升图像区域建模和文本短语的建模。算法团队把 mask 的内容从随机的单个单词和单个物体变成了临近的多个物体和连续的多个单词，从而模型在预测 mask 的内容时需要更大幅度地参考上下文信息，而由于单模态语义信息可能损失较大，模型需要更多地关注另一模态的信息从而实现更加准确的语言建模和图像建模。

在使用上述模型产出了多模态的特征之后，大模型可以迁移到具体的下游任务，如文本召回图片。相对于基线模型，多模态预训练模型在淘宝的电商场景中的文本召回图片的任务能够在召回指标 R@1、R@5 和 R@10 均能获得显著提升。而利用预训练模型产出的多模态特征也可以应用到推荐系统的商品召回任务中。算法团队使用丰富的电商商品文本和图片进行预训练，并对商品池中的每个商品产出多模态特征。这些高质量的多模态特征可以加入到推荐系统中的双流 i2i 召

AMiner

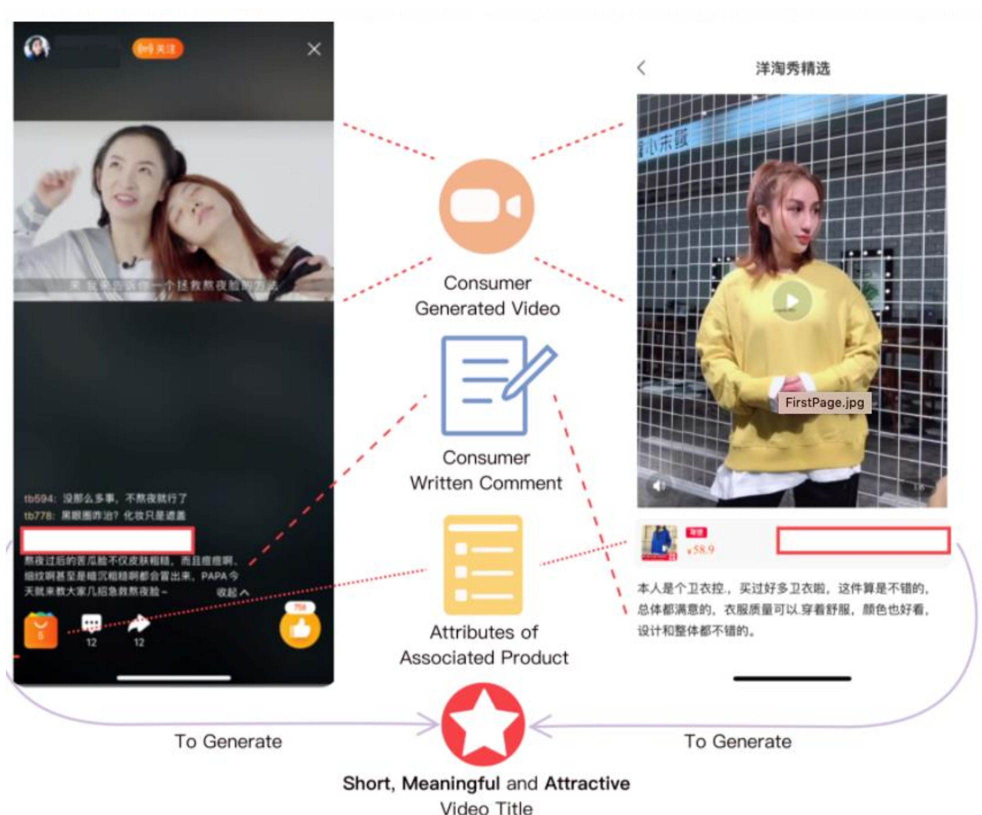
回模型中，让召回模型通过多模态特征建模用户行为。而由于多模态特征的作用，召回不会过分集中于只有大量用户行为的商品中，即常见的马太效应。目前预训练产出特征做的召回在线上也取得了喜人的表现，在点击率没有显著下降的情况下，曝光和点击的商品热度能够大幅降低，让更多位于商品分布中腰部的商品得到曝光。

（3） 基于用户交互的视觉智能

随着网络技术和大众传媒的发展，短视频和网络直播迈入了发展的快车道。无论是在抖音、快手等短视频 APP 中，还是在淘宝直播这种电商 APP 中，都产生了大量的 UGC 和 PGC 视频，视频成为了继文本、图像之后新兴的大众信息传媒载体。其中短视频时间短，内容丰富，或搞笑疗愈，或传授生活小知识，迅速收割了网民的碎片时间甚至整块休闲时间。同时，在电子商务领域，短视频在商品种草、商品展示、商品售后教程、视频买家秀等各个维度展示出了传统图文无法比肩的优势，在视频的数量和质量上得到了迅速提升。

电子商务领域的视频推荐对于获取新客户有着重要作用。例如，许多消费者会在商品评论区上传视频已分享他们独特的购物体验，这些独特的商品呈现方式或独特的商品使用方法可能会吸引潜在买家购买相同或类似的产品。相比于卖家秀视频（如广告），买家秀视频具有数量大和个性化强的优势，因此，将买家秀视频推荐给潜在感兴趣的消费者可以增强电子商务场景下视频推荐的有效性。下图展示了买家秀视频推荐的两个实际应用场景。

AMiner



AMiner

图 61 买家秀视频推荐的两个实际应用场景

买家秀视频在视频推荐过程缺少高质的视频标题。通常，视频推荐会给视频配上一个精炼的标题，该标题可以让目标用户在观看视频前对视频主要亮点、特色有一个初步把握，从而产生观看兴趣。虽然在消费者在商品评论区上传视频的过程中也会写有描述性的评论，但我们分析发现，这些评论经常会强调物流、客服和亲友评价等与商品本身外观和使用方法不相关的购物体验，不适宜作为推荐场景下的视频标题。

在这种应用需求下，我们设计了买家秀视频标题自动生成模型，并从以下数据中抽取、关联和聚合有用的信息：1）买家秀视频本身。视频以动态图像的形式展现了买家偏好的商品视觉特征以及拍摄主题；2）买家撰写的评论。虽然评论有较多的噪声信息，不能直接作为视频标题呈现，但是其内容可能蕴含了买家对产品特点的偏好以及使用体验。3）关联商品的属性信息。来自商品评论区的买家秀视频会与一个特定的商品关联，我们提取了关联商品的属性信息，如中长款（连衣裙），作为输入。属性结构化的展现了商品的主要特点。下图展示了一

个实际的数据样例。

图 62 商品关联属性信息图

目前和该任务设定接近的研究工作是视频描述生成任务（Video Captioning）。在任务本身上，一般性的视频描述生成数据集中的描述通常会只涵盖主要物体检测（Object Detection）和一般性的活动识别（Activity Recognition），如“A man is playing basketball”，而买家秀视频标题生成任务的描述涉及 商品细节特点识别、商品整体外观识别、商品和背景交互作用的识别（如 "夏日海滩风情度假必备连衣裙"）和视频故事线主题识别（如“这样穿显瘦美丽”）。因此在任务本身上，买家秀视频标题生成提出了更多和实际应用紧密相关的挑战；在现有模型层面，大部分视频描述模型将该任务看作序列到序列建模任务，典型的是采用带有或不带有注意力机制（Attention）的循环神经网络（RNN）同时建模视频帧序列和输出的词序列。然而，这样的建模方式具有局限性，其一，他们只在视频帧级别建模了视频信息，这不利于商品部位及商品整体级别的特点识别，基于序列建模的

RNN 模型本身也较难建模具有特殊依赖关系的商品不同部位；其二，序列建模的 RNN 模型无法系统和有效的关联和聚合三种异构信息，即视频、视频评论和关联商品的属性信息。

为了解决这些问题，我们提出了一个新的建模框架，名为基于图结构建模的视频标题生成器（Graph based Video Title Generator, Gavotte）。具体而言，Gavotte 由两个子过程组成，即细粒度交互建模和故事线摘要建模。细粒度交互建模将三种输入表示成图结构，并利用全局-局部聚合模块探索三种输入中的细粒度信息在图内和图间的交互作用，以识别用户偏好的商品细节特点和商品整体外观。故事线摘要建模利用了帧级别信息，更关注序列依赖，采用 RNN 结构，用以识别商品和背景交互作用和视频故事线主题。

我们在淘宝服饰类视频和非服饰类视频数据上对 Gavotte 进行了实验验证，并在手机淘宝上进行了上线测试，这些结果证明了 Gavotte 框架的有效性。

接下来本文将介绍 Gavotte 的重要组成结构以及主要的实验和分析。

- **Gavotte: Graph based Video Title Generator**

✧ **图表示**

首先我们介绍一下 Gavotte 如何将三种输入表示成图结构。

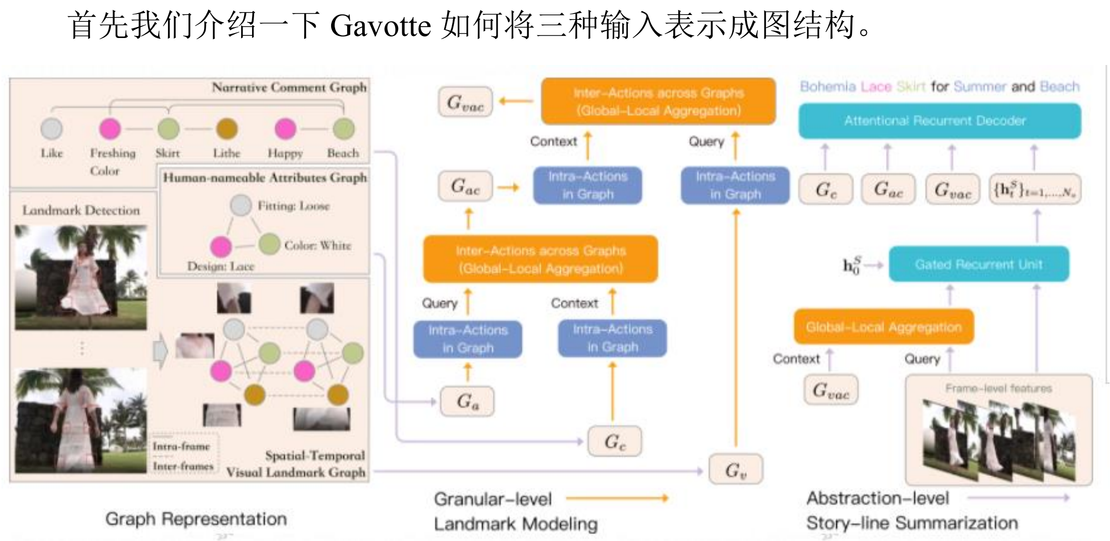


图 63 Gavotte 模型结构图

商品视频信息：我们首先使用了 Landmark Detection 技术（Liu et al. 2018）

检测衣服类商品的商品部位特征，每一帧的每一个部位都视为图节点。我们将同帧的商品部位节点进行全连接，不同帧相同部位节点进行全连接，这种连接方式一方面有利于我们捕捉同帧商品部位间的交互和商品整体风格，另一方面能够捕捉每一个商品部位跨越时间线的动态变化以及不同视点（pointview）下对商品部位的全面感知。为了增强空间-时序视觉商品部位图的时序性和部位特点，我们给每一个节点都加上了位置嵌入（Position Embedding）（Jonas et al. 2017）和类型嵌入（Type Embedding），即最终的节点表征为：

$$\mathbf{v}_{i,j}^{l} = \mathbf{l}_{i,j} + \mathbf{p}_i + \mathbf{r}_j$$

视频评论信息：我们将评论中的每一个词作为图节点，并将有语法依赖关系的节点进行连接，我们发现相比于时序关系，捕捉评论中和商品有关的语义信息对生成标题更有价值。

关联商品的属性信息：我们将每一个属性值（如 白色）作为图节点，并对所有节点进行全连接，属性间本身没有时序关系，使用图建模可以更好的探索属性间的特殊交互作用。

AMiner

✧ **细粒度交互建模**

如上图所示，细粒度交互建模对三个异构图的图内关系（Intra-Actions in Graph）和图间关系（Inter-Actions across Graphs）进行了建模。

✓ 图内关系建模注重和商品有关的细粒度特征的识别。我们采用了常用而有效的图神经网络作为图内关系建模的可训练框架结构，和最原始的图神经网络相比，本文对信息传播时的根节点和邻居节点进行了单独建模，并采用了自门控机制（Self Gating）。

$$\bar{\mathbf{v}}_i^g = \sigma(\mathbf{W}_r^g \mathbf{v}_i + \mathbf{W}_n^g \,\mathrm{MAX}(\{\mathbf{e}_{i,j} * \mathbf{v}_j, j \in \mathcal{N}(i)\}) + \mathbf{b}_n^g)$$
$$\bar{\mathbf{v}}_i = \bar{\mathbf{v}}_i^g * (\mathbf{W}_r^h \mathbf{v}_i + \mathbf{W}_n^h \,\mathrm{MAX}(\{\mathbf{e}_{i,j} * \mathbf{v}_j, j \in \mathcal{N}(i)\}) + \mathbf{b}_n^h)$$

✓ 图间关系建模注重异构图之间商品有关细粒度特征的关联和聚合。根据其实际的物理意义，我们将图间关系建模模块称为全局-局部聚合模块（Global-Local Aggregation，GLA）。GLA 包含全局门控访问（Global Gated Access）和局部注意（Local Attention）两个子模块。GLA 的输入为查询图（Query Graph）和

上下文图（Context Graph），输出的聚合图（Aggregated Graph）的结构与查询图一致。

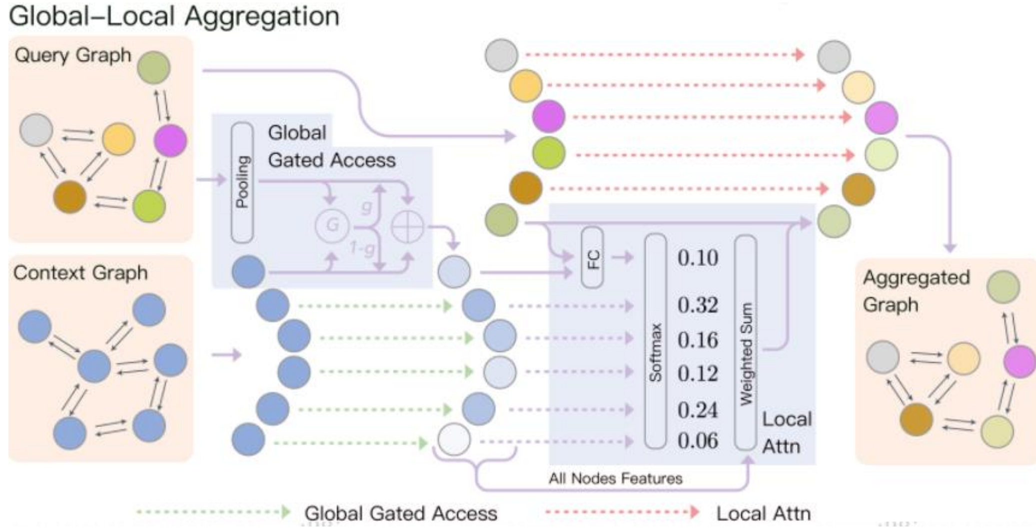


图 64 GLA 模型结构图

全局门控访问用来加强上下文图中与查询图全局相关的信息(可以视为初步筛选过程)，并抑制毫不相关的信息。

$$\mathbf{v}^{q*} = 1/N_q \sum_{i=1}^{N_q} \mathbf{v}_i^q$$

$$\mathbf{g}_j = \sigma(\mathbf{W}_g[\mathbf{v}^{q*}, \mathbf{v}_j^e] + \mathbf{b}_g)$$

$$\mathbf{v}_j^{eq} = (1 - \mathbf{g}_j) * (\mathbf{W}_q \mathbf{v}^{q*} + \mathbf{b}_q) + \mathbf{g}_j * (\mathbf{W}_e \mathbf{v}_j^e + \mathbf{b}_e)$$

✓ 局部注意力机制在局部层面（节点级别）筛选上下文图中和查询图节点相关的内容并聚合到该查询图节点上。

$$\mathbf{o}_{i,j} = \tanh(\mathbf{W}_o[\mathbf{v}_i^q, \mathbf{v}_j^{eq}] + \mathbf{b}_o)$$

$$\bar{\mathbf{o}}_{i,j} = \frac{\exp(\mathbf{W}_a \mathbf{o}_{i,j})}{\sum_k \exp(\mathbf{W}_a \mathbf{o}_{i,k})}$$

$$\mathbf{v}_i^{agg} = \mathbf{v}_i^q + \sum_{j=1}^{N_e} \bar{\mathbf{o}}_{i,j} * \mathbf{v}_j^{eq}$$

✧ **故事线摘要模块**

故事线摘要建模关注帧级别和视频级别的信息，因此采用帧特征和 RNN 来进行序列建模。首先我们利用全局-局部模块将细粒度信息和帧信息进行信息融

AMiner

合。这是因为我们发现视频帧建模（如商品-背景交互）和视频建模（如视频故事主题）和商品的细节特点息息相关。

$$\bar{\mathbf{X}} = GLA(\mathbf{X}, \mathbf{V}_{vac})$$

之后我们采用 RNN 对视频帧进行序列建模。

$$\begin{aligned} \tilde{\mathbf{x}}_t &= [\bar{\mathbf{x}}_t, \mathbf{x}_t] \\ \mathbf{h}_t^S &= \mathrm{GRU}(\mathbf{h}_{t-1}^S, \tilde{\mathbf{x}}_t) \end{aligned}$$

✧ 解码器

在解码器部分我们采用常用的注意力增强的 RNN 结构（Li et al. 2015），我们在每一步解码阶段都会关注细粒度图信息和帧信息。

$$\begin{aligned} &\mathbf{x}_m^D = [f_{la}(\mathbf{h}_m^D, \mathbf{V}_{vac}), f_{la}(\mathbf{h}_m^D, \mathbf{V}_{ac}), f_{la}(\mathbf{h}_m^D, \mathbf{V}_c), f_{la}(\mathbf{h}_m^D, \mathbf{H}^S)] \\ &\mathbf{h}_{m+1}^D = \mathrm{GRU}(\mathbf{h}_m^D, [\hat{\mathbf{c}}_m, \mathbf{x}_m^D]) \end{aligned}$$

✧ 学习目标

AMiner

我们采用了常用的交叉熵损失进行训练。

$$\mathcal{L}_{ce} = -\frac{1}{N_c}\sum_{m=1}^{N_c} \log f_P\left(c_m^R\right)$$

我们发现对于视频标题的生成，重复词的出现总是有损视频标题的整体吸引程度，因此我们参考了文本摘要（See et al. 2017）中常用的注意力覆盖损失，并提出了生成覆盖损失，用以惩罚和抑制重复词的生成。

$$P_m^O = softmax(\mathbf{W}_p \mathbf{h}_m^D + \mathbf{b}_p)$$

$$\varsigma_m = \sum_{m'=1}^{m-1} P_{m'}^O$$

$$\mathcal{L}_{gc} = \sum_m \sum_i \min\left(\varsigma_{m,i}, P_{m,i}^O\right)$$

✧ 实验和分析

我们摘选了主要的实验结果

Method	T-VTD				Other Categories			
	BLEU-4	METEOR	ROUGE_L	CIDEr	BLEU-4	METEOR	ROUGE_L	CIDEr
M-MPLSTM	0.95	10.43	15.47	32.65	1.17	7.58	13.12	25.75
M-S2VT	1.52	11.64	18.03	41.74	1.70	9.49	17.26	39.35
M-HRNE	1.70	12.11	18.85	44.53	1.66	9.40	17.10	39.36
M-SALSTM	2.01	12.95	19.81	48.63	**2.33**	10.65	18.42	47.01
M-RecNet	2.00	12.87	20.33	50.01	2.24	10.62	19.18	47.81
M-LiveBot	1.86	12.65	19.56	47.02	1.86	9.89	16.93	40.81
Gavotte	**2.28**	**13.58**	**21.33**	**54.14**	2.27	**10.99**	**19.38**	**49.13**

图 65 部分实验结果示例图

可以看到，我们的模型在两个淘宝商品数据集（T-VTD 服饰类商品数据，和其他商品类别数据）上去得了最佳的效果。我们的模型在 T-VTD 上提升显著，在其他商品类别数据上提升会相对弱一些，但这是合理的结果，因为服饰类商品有良好的商品部位定义，我们在其他类别上采用了近似的方法（细节可关注论文）。



图 66 M-Recnet 和 M-LiveBot 模型的结果对比图

该图展示了 Gavotte 和其他两个典型模型（基于 RNN 的 SOTA 模型 M-Recnet （Wang et al. 2018），和基于 Transformer 的 SOTA 模型 M-Livebot （Ma et al. 2019））的生成案例对比。具体而言，第一个案例中 M-Recnet 生成标题有意义信息较少，第二个案例中 M-LiveBot 生成标题出现未完成和损坏的情况，Gavotte 则生成了网络热词（如出街、抢镜），生成句子更为流畅，更有吸引力。同时，Gavotte 可以识别商品细节级别信息（如破洞），商品级别信息（如牛仔裤），商品与背景交互信息（如出街抢镜）和视频级别故事情节信息（如 这样穿）。

4.2 其他应用场景

4.2.1 行业应用

（1） 智慧城市建设

深兰科技在智慧城市的发展中强调认知智能，他们自主研发的具有全面自主知识产权的“智能交通协管员”以缓解城市交通繁忙，电子协警不会疲劳也不会有

情绪，语音系统人性化。在具体实践中，深兰科技把相关因素都抽取出来后，构建领域常识知识库，通过这些 Graph 来支撑这些规则，支撑这些方法有很多方法，比如 Scene Graph，还有视觉常识推理，以及基于机器推理的 MAC 网络。在这些有益的基础探索之上，构建了一些智慧城市智能城管系统，比如路政智能系统、智能交通管理系统、交通和城市管理系统，所有这些构建了深兰科技的平台，平台赋能这几大板块。

深兰科学院资深研究员曹东说新一代人工智能应该是可解释、小数据、大任务，而不是现在的大数据、小任务。在这个探索过程中，可以实现推理和认知，可以实现从知识图谱到事理图谱再到认知图谱的飞跃，认知图谱本质上也是知识图谱，无外乎是具有自学习、迭代、动态的过程，而不是像知识图谱那样只是一个静态的关系。为此，AI 要实现类似人类的能力，类比起来，实现推理，实现认知，它应当具备组合泛化和结构化的能力。组合泛化就是从已知构建块构建新的理论、预测和行为。

（2） 司法行业

AMiner

司法知识图谱将法律文书信息以结构化形式呈现，通过抽取案件文书中相关信息，判断案件繁简程度，进行繁简分流。这能帮助从业人员快速地在线检索相关的法务内容，提高法院审判工作质量和效率。将企业内部数据整合，打通数据孤岛，构建企业知识图谱，可以把人员、项目、产品等关联起来。基于知识图谱的语义搜索能力，能使搜索结果更准、更全。基于图谱的个性化推荐系统，可以实现知识精准传播和主动传播，让知识找人。这些应用都促进了沉淀知识的有效使用，极大地提高了工作效率。

（3） 金融行业

银行利用知识图谱可以进行借贷反欺诈。把与借款人相关的所有的数据源打通，并构建包含多数据源的知识图谱，把借款人的消费记录、行为记录、关系信息、线上日志信息等整合到反欺诈知识图谱里，从而进行分析和预测，可以挖掘识别出欺诈案件，如身份造假、团体欺诈、代办包装等。知识图谱还可以进行合同风险识别。通过自动提取文档内容的关键信息，协助企业工作人员完成内容一

致性检查等工作，自动生成对比结果报告。知识图谱能够自动识别数十种常见内置风险，根据不同行业特性支持个性化定制风险审核项。目前，知识图谱已经广泛应用于金融、制造、通信、法律、审计、政府等多种文字密集型行业，有效帮助识别风险。

（4） 安防行业

公安大数据应用的三大痛点：海量异构数据处理难、数据缺乏关联性、缺乏全警种智能应用。在公共安全领域，公安多年积累的实战经验与技术算法如何相互转换，是最大的行业难点，也是知识图谱主要解决的问题。公安知识图谱通过数据采集、数据处理、数据库重构、知识转化和实战应用五个步骤，运用分布式存储、关联算法、语义推理等技术，以及大量地公安专家团队与技术人员配合，来实现技术与业务的深度融合。目前，国内能够在“十亿数据节点，百亿关联边”的图谱结构下，实现秒级查询的公安知识图谱应用是较为优秀的产品标准，在国务院和公安部大力推动下，未来将有更大规模，更高数据集的知识图谱应用落地。

公安知识图谱建设起步较晚，2015 年有少部分大数据分析公司开始尝试建设，初期以小而具体的工具软件形式出现，数据关联更多是通过人工打标签实现，效率低且价值有限，在全国范围的应用率仅为 10%以下。通过技术发展和落地实践，以及企业内部技术标准统一和服务工具自动化迭代，自 2017 年开始市场认知发生变化，随着 2018 年公安部出台相关指导意见，各级公安系统对公安知识图谱的认识逐渐清晰、统一，更多基础层平台化的大型项目开始投入建设，已有项目维护和应用持续进行，2018 年工具型公安知识图谱应用率为 30%，平台型建设率达到 10%，艾瑞将持续看好其发展。

4.2.2 技术应用

（1） 大数据的精准与精细分析

如今，越来越多的行业或者企业积累了规模可观的大数据。但是这些数据并未发挥出应有的价值，很多大数据还需要消耗大量的运维成本。大数据非但没有创造价值，在很多情况下还成为了一笔负资产。这一现象的根本原因在于，当前的机器缺乏诸如知识图谱这样的背景知识，机器理解大数据的手段有限，限制了

AMiner

大数据的精准与精细分析，从而大大降低了大数据的潜在价值。举个亲身体验的例子，在娱乐圈王宝强离婚案刚刚开始的时候，新浪微博的热搜前三位分别是“王宝强离婚”、“王宝宝离婚”和“宝强离婚”。也就是说，当时的微博平台还没有能力将这三件事自动归类到一件事，不知道这三件事其实说的是一件事。机器在统计事件热度的时候就分开统计了，这就是因为当时机器缺乏背景知识，不知道王宝强又称为“王宝宝”或“宝强”，所以没有办法做到大数据的精准分析。事实上，舆情分析、互联网的商业洞察，还有军事情报分析和商业情报分析都需要大数据的精准分析，而这种精准分析就必须要有强大的背景知识支撑。除了大数据的精准分析，数据分析领域另一个重要趋势：精细分析，也对知识图谱和认知智能提出了诉求。比如很多汽车制造厂商都希望实现个性化制造。个性化制造希望从互联网上搜集用户对汽车的评价与反馈，并以此为据实现汽车的按需与个性化定制。为了实现个性化定制，厂商不仅需要知道消费者对汽车的褒贬态度，还需要进一步了解消费者不满意的细节之处，以及消费者希望如何改进，甚至用户提及了哪些竞争品牌。显然面向互联网数据的精细化数据分析必需要求机器具备关于汽车评价的背景知识（比如汽车的车型、车饰、动力、能耗等等）。因此，大数据的精准和精细化分析需要智能化的技术支撑。

AMiner

（2） 智慧搜索

下一代智慧搜索对机器认知智能提出了需求。智慧搜索体现在很多方面。首先，体现在搜索意图的精准理解方面。比如在淘宝上搜索“iPad 充电器”，用户的意图显然是要搜索一个充电器，而不是一个 iPad，这个时候淘宝应该反馈给用户若干个充电器以供选择，而不是 iPad。再比如在 Google 上搜索“toys kids”或者“kids toys”，不管搜索这两个中的哪一个，用户的意图都是在搜索给孩子的玩具，而不是玩玩具的小孩，因为一般不会有人用搜索引擎搜孩子。“toys kids”和“kid’s toys”中两个词都是名词，要辨别出哪一个是核心词，哪一个是修饰词，在缺乏上下文的短文本上，仍然是个具有挑战性的难题。其次，搜索的对象越来越复杂多元化。以前搜索的对象以文本为主，现在大家希望能搜索图片和声音，甚至还能搜代码，搜视频，搜设计素材等等，要求一切皆可搜索。第三、搜索的粒度也越来越多元化。现在的搜索不仅能做篇章级的搜索，还希望能做到段落级、语句级、

词汇级的搜索。尤其是在传统知识管理领域，这个趋势已经非常明显。传统的知识管理大都只能做到文档级搜索，这种粗粒度的知识管理已经难以满足实际应用中细粒度的知识获取需求。最后，是跨媒体的协同搜索。传统搜索以面向单质单源数据的搜索居多。比如针对文本搜索难以借力视频、图片信息，针对图片的搜索主要还是利用图片自身的信息，对于大量文本信息利用率还不高。最近的趋势是跨媒体的协同搜索。比如前几年，明星王珞丹在微博上晒了张自家小区的照片，然后就有好事者根据她的微博社交网络、百度地图、微博文本与图片信息等多个渠道多种媒体的信息，通过联合检索准确推断出其所在小区位置。所以，未来的趋势是一切皆可搜索，并且搜索必达。

（3） 智能推荐

智能推荐表现在很多方面。首先是场景化推荐。比如用户在淘宝上搜"沙滩裤"、"沙滩鞋"，可以推测这个用户很有可能要去沙滩度假。那么平台是否能推荐"泳衣"、"防晒霜"之类的沙滩度假常用物品呢？事实上，任何搜索关键字背后，购物篮里的任何一件商品背后都体现着特定的消费意图，很有可能对应到特定的消费场景。建立场景图谱，实现基于场景图谱的精准推荐，对于电商推荐而言至关重要。第二、任务型推荐。很多搜索背后的动机是完成特定任务。比如用户购买了"羊肉卷"、"牛肉卷"、"菠菜"、"火锅底料"，那么用户很有可能是要做一顿火锅，这种情况下，系统推荐火锅调料、火锅电磁炉，用户很有可能买单。第三、冷启动下的推荐。冷启动阶段的推荐一直是传统基于统计行为的推荐方法难以有效解决的问题。利用外部知识，特别是关于用户与物品的知识指引冷启动阶段的匹配与推荐，是有可能让系统尽快渡过这个阶段的。第四、跨领域的推荐。当阿里刚刚入股新浪时，我们在设想是否能将淘宝的商品推荐给微博的用户。比如，如果一个微博用户经常晒九寨沟、黄山、泰山的照片，那么为这位用户推荐一些淘宝的登山装备准没错。这是典型的跨领域推荐，微博是一个媒体平台，淘宝是一个电商平台。他们的语言体系、用户行为完全不同，实现这种跨领域推荐显然商业价值巨大，但却需要跨越巨大的语义鸿沟。如果能有效利用知识图谱这类背景知识，不同平台之间的这种语义鸿沟是有可能被跨越的。比如百科知识图谱告诉我们九寨沟是个风景名胜，是个山区，山区旅游需要登山装备，登山装备包括

AMiner

登山杖、登山鞋等等，从而就可以实现跨领域推荐。第五、知识型的内容推荐。在淘宝上搜索“三段奶粉”，能否推荐“婴儿水杯”，同时我们是否能推荐用户一些喝三段奶粉的婴儿每天的需水量是多少，如何饮用等知识。这些知识的推荐，将显著增强用户对于推荐内容的信任与接受程度。消费背后的内容与知识需求将成为推荐的重要考虑因素。所以未来的推荐趋势就是精准感知任务与场景，想用户之未想。推荐技术演进的重要趋势是从单纯基于行为的推荐过渡到行为与语义融合的推荐。换言之，基于知识的推荐，将逐步成为未来推荐技术的主流。

（4） 智能解释

2017 年年底的时候，微信上流传 Google17 年最流行的搜索关键字是“how”，这说明人们希望 Google 平台能做解释。类似于“如何做蛋炒饭”，“怎么来北理工”等这类问题在搜索引擎上出现次数日益增多，这些问题都在考验机器的解释水平。一个更有意思的例子是，当我们在 Google 上搜索“Donald Trump”相关的问题时，你会发现 Google 会自动提示询问“为什么特朗普的老婆嫁给他”而不再是“特朗普老婆是谁”这类简单事实问题。“why”和“how”问题在现实应用中越来越多。这一趋势实际上体现了人们的一个普遍诉求，那就是希望智能系统具备可解释性。所以可解释性将是智能系统一个非常重要的体现，也是人们对智能系统的普遍期望。

可解释性决定了 AI 系统的决策结果能否被人类采信。可解释性成为了很多领域（金融、医疗、司法等）中阻碍 AI 系统落地应用的最后一公里。比如在金融领域的智能投资决策，即便 AI 决策的准确超过 90%，但是如果系统不能给出作出决策的理由，投资经理或者用户恐怕也是十分犹豫的。再比如在医疗领域，即便系统判断疾病的准确率在 95%以上，但是如果系统只是告诉病患得了什么病或者开了一个药方，却不能解释为什么作出这类判断的话，病人是不会为此买单的。

智能系统的可解释性体现在很多具体任务中，包括解释过程、解释结果、解释关系、解释事实。事实上，可解释人工智能最近受到了越来越多的关注。在学术界，机器学习，特别是深度学习的黑盒特性，日益成为学习模型实际应用的主要障碍之一。越多越多的学术研究项目旨在打开深度学习的黑盒。美国军方也有项目在尝试解释机器的学习过程。我个人也曾做过“基于知识图谱的可解释人工

AMiner

智能”的研究与思考，旨在强调知识图谱在可解释性方面的重要作用。

（5） 自然人机交互

智能系统另外一个非常重要的表现方式就是自然人机交互。人机交互将会变得越来越自然，越来越简单。越自然简单的交互方式越依赖强大的机器智能水平。自然人机交互包括自然语言问答、对话、体感交互、表情交互等等。特别是自然语言交互的实现需要机器能够理解人类的自然语言。会话式（Conversational UI）、问答式（QA）交互将逐步代替传统的关键字搜索式交互。对话式交互还有一个非常重要的趋势就是一切皆可问答。我们的 BOTs（对话机器人）将代替我们阅读文章、新闻，浏览图谱、视频，甚至代替我们看电影、电视剧，然后回答我们所关心的任何问题。自然人机交互的实现显然需要机器的较高认知智能水平，以及机器具备强大的背景知识。

人工智能的重要发展方向，是人机混合智能。强人工智能、通用人工智能，及类人人工智能，实现还相对较遥远，当前相对实现的途径就是人机混合智能。人机混合智能主要研究如何在人、机及环境系统之间，实现最优的智能匹配，人的智能加上机器的智能，涉及人机环境系统的整体设计，及其优化等方面的研究，研究目的包括可靠、高效和舒适等几个方面。为了实现以上目标，需要解决人的意向性和机器的形式化问题。所谓意向性，就是意识的指向。机器难以处理涉及到灵活、可改变的，甚至带有矛盾性的事物，但是机器的长处在于它不疲劳、擅计算，并且能够准确及时的，处理形式化、符号化的东西，而这是人所不擅长的。所以，如何把机器的长处，和人的这种优点，充分的结合在一起，这是一个很重要的命题，也是人机混合智能的一个命脉。

未来的人机交互及人工智能系统，有明确的发展方向，它包含四个方面：主动的推荐、自主的学习、自然的进化、自身的免疫。在这四个方面中自主性是非常重要的一个概念，应该包括以下几方面：1）自主应具有记忆的功能，而不是存储，记忆是灵活的，能够通过相关无关的事物产生直觉，而存储则无法出现直觉，它只是符合逻辑的东西；2）自主应具有选择性，选择性是单向性的，即 A 选择 B；3）自主应具有匹配性，匹配和选择最大的区别，就是匹配是双向性的，A 可以选择 B，B 也可以选择 A；4）自主应可以控制，没有控制和反馈，自主很

AMiner

难建立起来。未来的人机交互智能系统需要在输入、处理、输出和反馈等信息流动过程中具有认知智能，至少是人机环境系统的自主耦合[183]。

AMiner

5　趋势篇



5 趋势篇

本节根据附录 1 展示的与认知图谱领域研究相关的关键词列表，依托 AMiner 平台技术分析系统（http://trend.aminer.cn），搜索相关研究论文并进行深入挖掘，探索分析了认知图谱技术的研究发展趋势和创新热点等，为科技工作者了解认知图谱领域相关技术的研究历史和现状、快速识别前沿热点研究问题提供信息窗口。

5.1 技术研究发展趋势

技术研究发展趋势详细展示了领域技术的开始研究时间、历年的研究热度变化情况，对于学科领域的布局和发展具有重要的意义和价值。认知图谱领域部分技术关键词，包括：Knowledge Base（知识库）、Semantic Web（语义 Web）、Natural Language Processing（自然语言处理）、World Wide Web（万维网）、Data Mining（数据挖掘）、Machine Learning（机器学习）、Neural Networks（神经网络）、Knowledge Representation（知识表示）、Ontology（本体）、Expert System（专家系统），的研究发展趋势如图 67 所示。图中每条色带表示一个研究话题，色带宽度表示该话题在当年的研究热度，与当年该话题的论文数量呈现正相关关系，不同色带的高低排序是由当年这些话题的研究热度决定。

从图中可以看出，这些话题的研究热度随着时间推进呈现增长趋势。尤其是近几年，知识库、语义 Web、自然语言处理的研究热度居高不下。在技术层面，机器学习和神经网络模型一直是该领域应用广泛的技术，尤其是近十年来，它们的研究热度出现明显的上升趋势。此外，从图中还可以看出，专家系统在 20 世纪 70 年代出现以来，研究热度一开始显著提升，后来出现下滑趋势。这是因为专家系统模拟人类专家的知识和经验解决特定领域的问题，成效显著，推动人工智能走入应用发展的新高潮，进而提升了其研究热度。但是后来经过实践应用，人们发现专家系统存在应用领域狭窄、缺乏常识性知识、知识获取困难、推理方法单一、缺乏分布式功能和难以与现有数据库兼容等问题，导致专家系统的研究热度逐渐下降。

AMiner

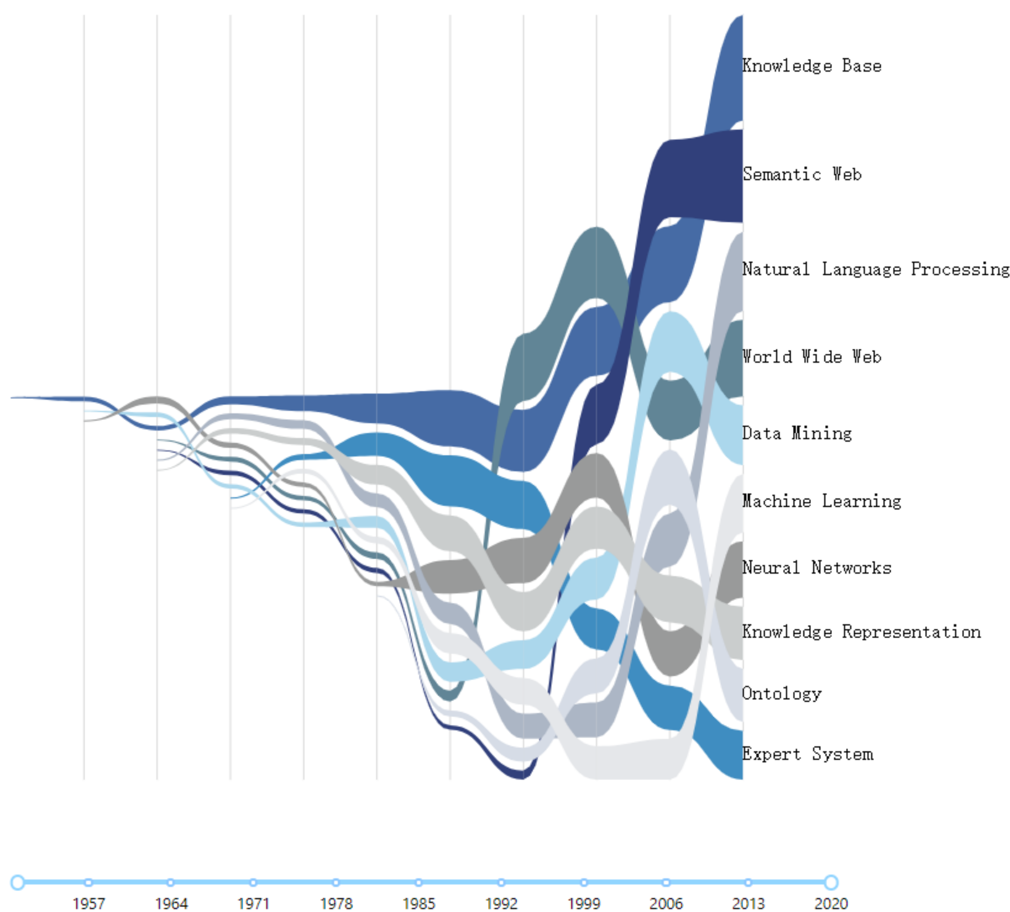


图 67 认知图谱领域的技术研究发展趋势

5.2 技术研究创新热点

图 68 展示了认知图谱领域研究热度较高的话题分布。其中每个关键词的大小表示研究热度高低，与研究的论文数量成正比。从中可以看出，该领域研究热度较高的是 Knowledge Base、World Wide Web、Semantic Web、Natural Language Processing、Ontology 等。将 h-index 指标作为筛选条件，从 AMiner 平台上获取了以上几个研究热度较高的领域的代表性学者信息，包括 Knowledge Base 研究领域的代表性学者是马克斯普朗克信息学研究所的 Gerhard Weikum 主任，他获得了清华大学 AMiner 团队发布的“AI 2000 人工智能全球最具影响力学者榜单[2]”（简称 AI 2000）的信息检索与推荐领域的最具影响力学者奖优秀奖。World Wide

[2] AI 2000 旨在通过 AMiner 学术数据在全球范围内遴选 2000 位人工智能学科最有影响力、最具活力的顶级学者。AI 2000 涵盖人工智能学科 20 个子领域，具体遴选方法为每个子领域每年选出 10 名获奖者，未来 10 年共产生 2000 名；每年遴选时，参考过去 10 年该领域最有影响力的期刊和会议所发表论文的引用情况，排名前 10 的学者当选该领域当年【AI 2000 最具影响力学者奖】。



图 68 认知图谱领域的技术创新热点词云图

AMiner

Web 领域的代表性学者是康奈尔大学计算机科学系的 Jon M. Kleinberg 教授，他也获得了 AI 2000 榜单的信息检索与推荐领域的最具影响力学者奖优秀奖。Semantic Web 领域的代表性学者是美国南卡罗来纳大学的 Amit P. Sheth 教授，他获得了 AI 2000 榜单的知识工程领域的最具影响力学者奖优秀奖。Natural Language Processing 领域的代表性学者是斯坦福大学的 Christopher D. Manning 教授，他获得了 AI 2000 榜单的自然语言处理领域的最具影响力学者奖优秀奖。Ontology 领域的代表性学者是美国哈佛大学的 Ian Horrocks 教授，他也获得了 AI 2000 榜单的知识工程领域的最具影响力学者奖优秀奖。根据 AMiner 平台的不完全统计，表 12 展示了这些学者的学术指标。

表 12 认知图谱研究热点子领域的代表性学者的学术指标统计

学者姓名	论文发表量	论文总被引频次	H-index	G-index	Sociability	Diversity	Activity
Gerhard Weikum	884	48100	93	208	7	4	221
Jon M. Kleinberg	421	90867	118	301	6	4	99
Amit P.	832	44506	103	200	7	4	100

Sheth							
Christopher D. Manning	464	126766	121	355	6	3	62
Ian Horrocks	419	54538	98	232	7	3	112

5.3 中国专利数据情况

本报告根据"知识库（Knowledge base）、语义网络（Semantic network）、知识图谱（Knowledge graph）、认知智能（Cognitive intelligence）、知识图谱推理（Knowledge graph inference/ reasoning）、图推理（Graph reasoning）、自然语言生成（Natural language generation）、文本生成（Text generation）、自动文摘（Text summarization）、图表示（Graph representation）"等中英文关键词检索到中国在 2010-2019 年申请和授权的 52242 篇专利。

图 69 展示了中国在 2010-2019 年间的历年专利数量的分布情况。从图中可以看出，中国近 10 年的专利数量整体上呈现逐年递增趋势，尤其在 2011 年的增长率最高（约 90.33%）。

AMiner

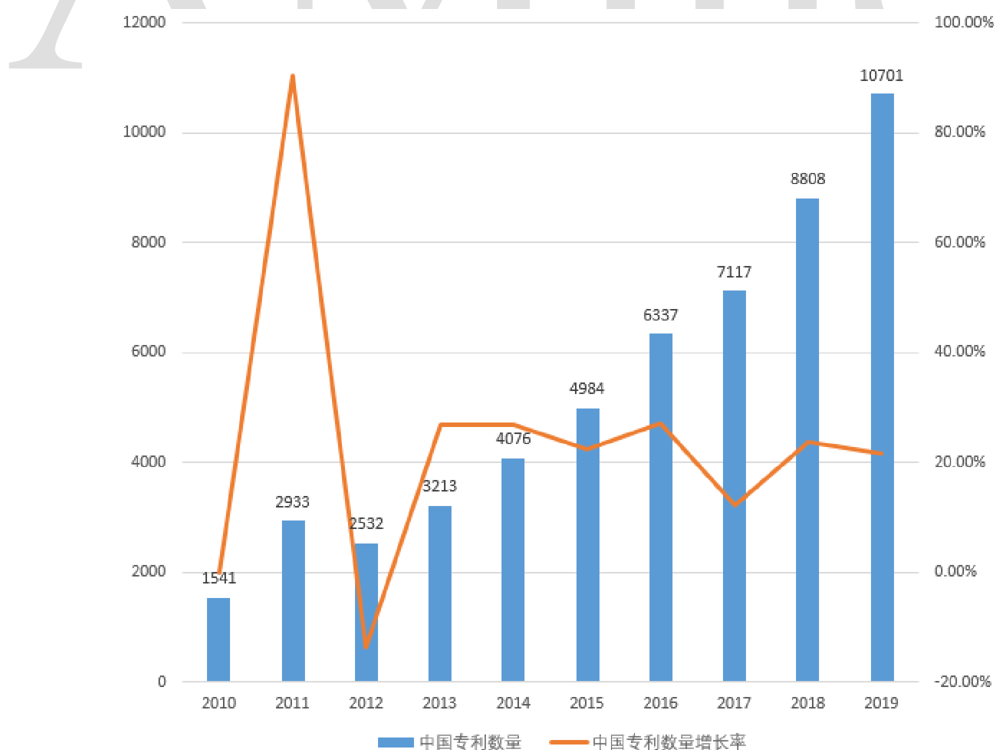


图 69 中国历年的专利数量分布（2010-2019 年）

图 70 展示了中国在 2010-2019 年间专利数量最多的前 10 个机构。这 10 个

机构中包含 6 所公司企业和 4 所高校，主要分布在北京（4 所）、广东（3 所）、陕西（1 所）、浙江（2 所）等地区。

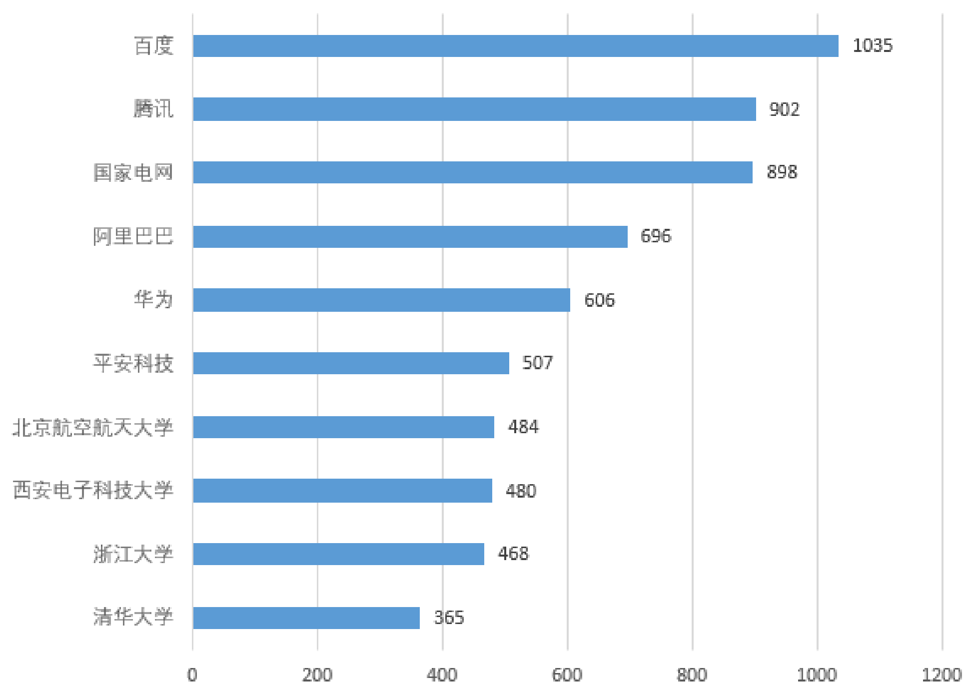


图 70 2010-2019 年中国专利数量 TOP 10 机构

AMiner

5.4 国家自然科学基金支持情况

本报告根据附录 1 中关键词（包含中文翻译），结合第 2 章介绍的技术领域，从 AMiner 数据库中查找到 522 个国家自然科学基金在 2010 至 2020 年支持的相关项目（包含未结题的项目），如表 13 所示。522 个项目主要分为 11 类，其中面上项目和青年科学基金项目占据绝大多数，占比约 78.1%。

表 13 认知图谱相关领域国家自然科学基金项目分类情况（2010-2020 年）

项目类别	项目数量	数量占比
面上项目	216	41.4%
青年科学基金项目	191	36.6%
地区科学基金项目	39	7.5%
联合基金项目	29	5.6%
重大研究计划	18	3.4%
专项基金项目	10	1.9%
应急管理项目	7	1.3%
重点项目	6	1.1%
国际(地区)合作与交流项目	3	0.6%
国家基础科学人才培养基金	2	0.4%

优秀青年科学基金项目	1	0.2%

图 71 展示了认知图谱领域国家自然科学基金支持项目在 2010 年到 2020 年的分布情况。从中可以看出，每年的支持项目数量相差不大。从整体分布来看，支持项目数量开始呈现上升趋势，2014 年达到顶峰，以后几年数量逐年下降，直到 2020 年数量开始上升。这种现象一方面原因是和谷歌公司在 2012 年提出了 Knowledge Graph 概念，并研发该项目用于提升搜索引擎质量，极大带动了学术界和工业界的研究热情。但由于知识抽取困难限制了知识图谱的规模，使得知识图谱发挥的作用有限，人们降低了对知识图谱的研究热情。最近两年，具有认知能力的人工智能需要有数据和知识作为支撑等观点提出，侧面推动了知识图谱的研究，尤其是常识知识图谱的研究。

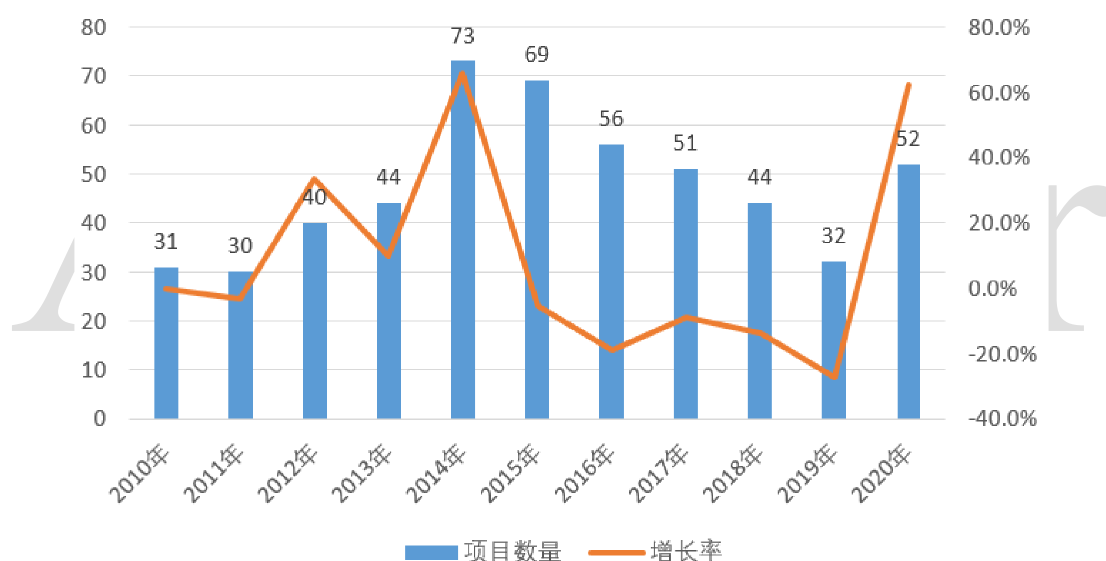


图 71 认知图谱相关领域国家自然科学基金项目支持历年分布情况

图 72 展示了获得国家自然科学基金支持项目（认知图谱相关领域）数量 TOP 10（实际为 11 个）的中国机构列表。这些机构的地区分布比较均匀，主要分布在东部地区的北京市（3 个）、上海市（1 个）、浙江省（1 个）、江苏省（1 个），中部地区的湖北省（1 个），西南地区的广西省（1 个），以及东北地区的吉林省（1 个）、辽宁省（2 个）。其中，桂林电子科技大学的项目类别主要是地区科学基金项目（8 个）、青年科学基金项目（2 个）和联合基金项目（1 个）。

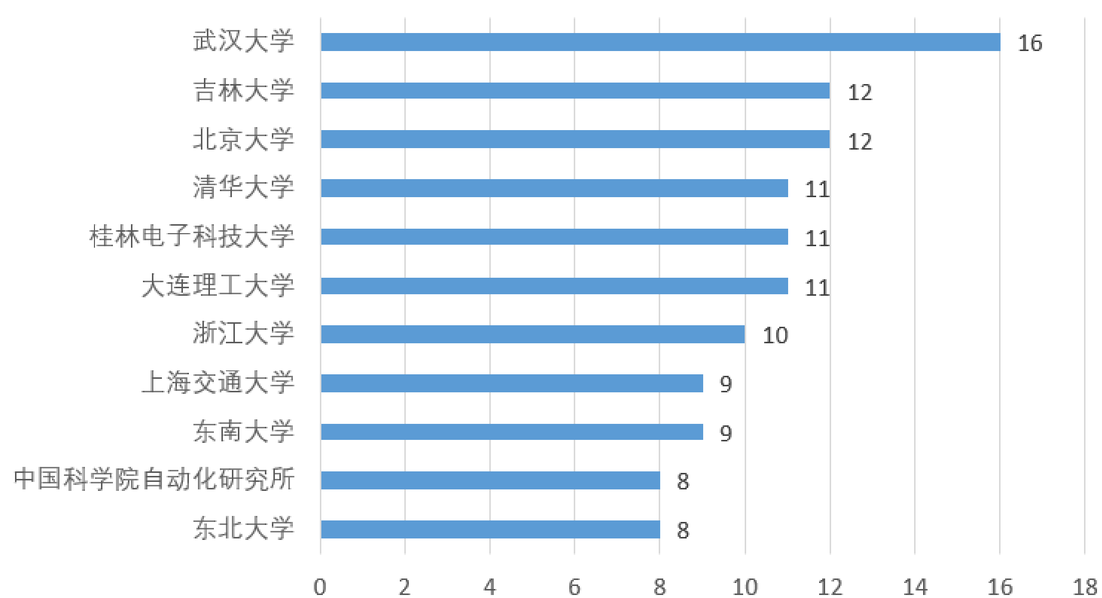


图 72 认知图谱相关领域国家自然科学基金项目支持数量 TOP 10 机构统计

AMiner

6　总结与展望



6 总结与展望

1956 年，人工智能概念提出以后，经过六十多年的发展，人工智能在计算能力、大数据和深度学习的支撑下取得显著成果。近年来，人工智能技术迅速普及应用，走进人们的日常生活，比如手机上的语音识别、人脸识别、机器翻译等。仔细观察就会发现，这些技术主要集中在感知层面，即用人工智能模拟人的感知能力，如听觉、视觉等。人类社会已经进入智能时代，智能时代的社会发展催生了大量的智能化应用，智能化应用对机器的认知智能化水平提出了前所未有的要求，而现有的知识图谱技术难以支撑机器去解决推理、规划、联想、创作等复杂的认知智能化任务。

2015 年，张钹院士提出第三代人工智能体系的雏形，并于 2018 年底，第三代人工智能的理论框架体系正式公开，其核心思想为：（1）建立可解释、鲁棒性的人工智能理论和方法；（2）发展安全、可靠、可信及可扩展的人工智能技术；（3）推动人工智能创新应用。第三代人工智能的理念在国内外获得广泛影响力。该框架具体实施的路线图是：（1）与脑科学融合，发展脑启发的人工智能理论；（2）数据与知识融合的人工智能理论与方法[184]。2017 年 DARPA（美国国防高级研究计划局）发起 XAI 项目（explainable AI，可解释人工智能），从可解释的机器学习系统、人机交互技术以及可解释的心理学理论三个方面，全面开展可解释性 AI 系统的研究。2020 年 1 月，阿里达摩院发布了“2020 十大科技趋势”，报告指出，人工智能未来将实现两大突破：认知智能与机器间大规模协作。这两大关键技术一方面可以将人工智能的能力从基础的“听说看”拓展到逻辑推理、思考决策，进而使人工智能深度介入社会生产生活；另一方面可以将智能体大规模自主组织起来，不再需要大量的人工操作，进而使人工智能真正成为协调整个人类工作生活网络的核心系统。

如何实现认知智能，清华大学唐杰教授提出了结合知识图谱、认知推理、逻辑表达等关键技术的认知图谱是一种有力的支撑手段。认知图谱，涉及语义理解、知识表达、联想推理、智能问答、自主学习等。认知图谱的发展将使大量繁琐却重要的工作变得更加高效精准，也更人性化。为了帮助读者对认知图谱相关技术有个整体认识和了解，本报告围绕认知图谱，以及知识图谱、认知推理、逻辑表

AMiner

达等关键技术，从概念背景、发展历程、实现技术、应用、问题与与挑战、未来研究方向等角度展开介绍，为读者了解认知图谱相关领域的发展动向、基础及应用研究的代表性成果提供了信息窗口。

本报告依托 AMiner 平台，对技术领先的国家、机构和学者，以及认知图谱相关领域的技术发展趋势、技术创新热点等进行了分析，旨在为国家和机构在该领域学科的研究发展和布局提供数据支撑和参考。通过对认知图谱相关领域的论文和学者数据进行统计分析，总结了中国高水平学者队伍建设过程中的问题并给予了相应的对策建议。

将人工智能真正转变成现实的生产力，离不开认知智能的转化应用，而将认知智能推入发展的快车道，还有一段路要走。清华大学唐杰教授在“人工智能下一个十年”报告中提到，要真正实现知识和推理，其实还需要万亿级的常识知识库的支持。现在需要构建更大规模的常识知识图谱，来支撑深度学习的计算，这样才能真正实现具有认知智能的 AI。图灵奖得主 Manuel Blum（卡内基梅隆大学计算机科学教授，计算复杂性理论的主要奠基人之一）在“2020 北京智源大会”上分享了主题为“*Towards a Conscious AI: A Computer Architecture inspired by Neuroscience*”的报告，核心理念是构造一个新型的可用数学建模、可计算的机器认知模型，以此实现有意识、可解释的人工智能。

“神经+符号”系统因为背景知识无缝利用、数据噪音鲁棒性强、求解过程和结果容易理解和解释等优点，是实现新一代人工智能模型的理想模型，但是如何有机结合来达到神经和符号之间的平衡是实现模型价值最大化的关键。认知图谱是以图结构的方式描述客观世界中的事物及其关联关系。考虑到图神经网络在数据建模、信息传播、关系归纳偏置等任务上的优秀性能，将其与认知图谱结合，融合知识和数据，有助于提升人工智能技术的可解释性和可信性。基于认知图谱的知识组织结构，利用图神经网络进行知识推理，是实现认知推理的有效手段，未来需要重点关注如何同时捕捉结构信息和学习规则知识，以此提升认知图谱推理的性能。

近年来，在机器学习和深度学习技术的推动下，人工智能在感知计算层面取得了显著成果，但是如何将深度学习与大规模常识知识结合起来，实现鲁棒的、

AMiner

可解释的、有意识能决策的新一代人工智能还面临很大挑战。大数据环境下，大规模、形式多样、零散分布、动态变化、质量低下的数据特征给自然语言处理、知识工程等人工智能技术提出了新的挑战，不仅需要从感知角度学习数据的分布表示，还需要从认知角度解释数据的语义。研发融合常识知识图谱、认知推理和逻辑表达等核心技术的认知图谱将成为实现下一代人工智能技术突破的关键。实现以认知图谱作为底层数据支撑，具有推理、具有可解释性、具有认知的新一代人工智能，是人工智能下一个十年必须发展的重要方向，将预示下一波人工智能浪潮的兴起。

AMiner

参考文献

[1] 闽南网.阿里巴巴达摩院发布了 2020 十大科技趋势[EB/OL]. http://www.mnw.cn/keji/internet/2236878.html

[2] 搜狐.复旦大学肖仰华教授：知识图谱与认知智能[EB/OL]. https://www.sohu.com/a/227760822_500659

[3] CSDN.CNCC_2018----"认知图谱与知识推理".[EB/OL] https://blog.csdn.net/qq_38150441/article/details/83512669.

[4] 知乎.人工智能下一个十年[EB/OL]. https://zhuanlan.zhihu.com/p/126656200.

[5] 百度百科.认知图谱[EB/OL]. https://baike.baidu.com/item/%E8%AE%A4%E7%9F%A5%E5%9B%BE%E8%B0%B1/24409851?fr=aladdin

[6] Quillian, M. R. Semantic memory. M Minsky, Ed, Semantic Information Processing[J]. 1968: 216-270.

[7] Buchanan, B., Georgia Sutherland, and E. A. Feigenbaum. Heuristic DENDRAL: A program for generating explanatory hypotheses[J]. Organic Chemistry. 1969.

[8] Gruber, Thomas R. The role of common ontology in achieving sharable, reusable knowledge bases[J]. Kr 91. 1991: 601-602.

[9] Berners-Lee, Tim. Semantic web road map[J]. emse.fr. 1998.

[10] Google.Introducing the Knowledge Graph: things, not strings[EB/OL]. http://blog.google/products/search/introducing-knowledge-graph-things-not

[11] InfoQ.肖仰华：知识图谱下半场 - 机遇与挑战[EB/OL]. https://www.infoq.cn/article/UjqxsX2uCYa5Fe7jNDT2?utm_source=related_read&utm_medium=article

[12] Richens R H. Preprogramming for mechanical translation[J]. Mechanical Translation, 1956, 3(1): 20-25.

[13] Newell A, Shaw J C, Simon H A. Report on a general problem solving program[C]//IFIP congress. 1959, 256: 64.

[14] Computer-based medical consultations: MYCIN[M]. Elsevier, 2012.

[15] Cycorp.CYC[EB/OL].https://www.cyc.com/

[16] W3C.Resource Description Framework(RDF) Model and Syntax Specification[EB/OL].

AMiner

https://www.w3.org/TR/1999/REC-rdf-syntax-19990222/
[17] W3C.OWL Web Ontology Language Guide[EB/OL].https://www.w3.org/TR/owl-guide/
[18] W3C.W3C[EB/OL].https://www.w3.org/standards/semanticweb/
[19] Stokman F N , Vries P H D . Structuring knowledge in a graph[M]// Human-Computer Interaction. Springer Berlin Heidelberg, 1988.
[20] Dong X , Gabrilovich E , Heitz G , et al. Knowledge vault: a web-scale approach to probabilistic knowledge fusion[J]. 2014.
[21] Ji S , Pan S , Cambria E , et al. A Survey on Knowledge Graphs: Representation, Acquisition and Applications[J]. 2020.
[22] Wang, Haofen. Zhishi.me – weaving Chinese linking open data.[J].
[23] Oren Etzioni, Michael J. Cafarella, Doug Downey,et al. Web-scale information extraction in knowltAll (preliminary results)[C]// Proceedings of the 13th international conference on World Wide Web, WWW 2004, New York, NY, USA, May 17-20, 2004. 2004.
[24] Alexander Yates, Michele Banko, Matthew Broadhead,et al. TextRunner: Open Information Extraction on the Web[C]// Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics. Association for Computational Linguistics, 2007.
[25] Carlson A , Betteridge J , Kisiel B , et al. Toward an Architecture for Never-Ending Language Learning[C]// Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2010, Atlanta, Georgia, USA, July 11-15, 2010. AAAI Press, 2010.
[26] 李涓子, 侯磊. 知识图谱研究综述[J]. 山西大学学报:自然科学版, 2017(3).
[27] Wang Q , Mao Z , Wang B , et al. Knowledge Graph Embedding: A Survey of Approaches and Applications[J]. IEEE Transactions on Knowledge & Data Engineering, 2017, 29(12):2724-2743.
[28] Mikolov T, Yih W, Zweig G. Linguistic regularities in continuous space word representations[C]//Proceedings of the 2013 conference of the north american chapter of the association for computational linguistics: Human language technologies. 2013: 746-751.
[29] Bordes A, Usunier N, Garcia-Duran A, et al. Translating embeddings for modeling multi-relational data[C]//Advances in neural information processing systems. 2013: 2787-2795.
[30] Wang Z , Zhang J , Feng J , et al. Knowledge Graph Embedding by Translating on Hyperplanes[C]// Twenty-eighth Aaai Conference on Artificial Intelligence. AAAI Press, 2014.
[31] Shaozhi, Dai, Yanchun, et al. Learning Entity and Relation Embeddings with Entity Description for Knowledge Graph Completion[C]//.
[32] He S , Liu K , Ji G , et al. Learning to Represent Knowledge Graphs with Gaussian
AMiner

Embedding[C]// Acm International. ACM, 2015.

[33] Xiao H, Huang M, Zhu X. TransG: A generative model for knowledge graph embedding[C]// Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2016: 2316-2325.

[34] Nickel M , Tresp V , Kriegel H P . A Three-Way Model for Collective Learning on Multi-Relational Data[C]// Proceedings of the 28th International Conference on Machine Learning, ICML 2011, Bellevue, Washington, USA, June 28 - July 2, 2011. Omnipress, 2011.

[35] Rodolphe Jenatton, Nicolas Le Roux, Antoine Bordes,等. A latent factor model for highly multi-relational data[C]// Advances in Neural Information Processing Systems. Curran Associates Inc. 2012.

[36] Alberto García-Durán, Bordes A , Usunier N . Effective Blending of Two and Three-way Interactions for Modeling Multi-relational Data[C]// Joint European Conference on Machine Learning & Knowledge Discovery in Databases. Springer, Berlin, Heidelberg, 2014.

[37] Yang B , Yih W T , He X , et al. Embedding Entities and Relations for Learning and Inference in Knowledge Bases[J]. 2014.

[38] Nickel, Maximilian, Rosasco, Lorenzo, Poggio, Tomaso. Holographic Embeddings of Knowledge Graphs[J].

[39] Bordes A, Glorot X, Weston J, et al. A semantic matching energy function for learning with multi-relational data[J]. Machine Learning, 2014, 94(2): 233-259.

[40] 刘峤，李杨，段宏等.知识图谱构建技术综述[J]. 计算机研究与发展(3):582-600,共 19 页.

[41] Huffman S B. Learning information extraction patterns from examples[J]. 1996.

[42] Grishman R, Sundheim B. Message understanding conference-6: A brief history [C]. In COLING 1996 Volume 1: The 16th International Conference on Computational Linguistics. 1996.

[43] Yu L-C, Lee L-H, Tseng Y-H, et al. Overview of SIGHAN 2014 Bake-off for Chinese spelling check [C]. In Proceedings of The Third CIPS-SIGHAN Joint Conference on Chinese Language Processing. 2014: 126–132.

[44] Tjong Kim Sang E F, De Meulder F. Introduction to the CoNLL-2003 shared task: Language-independent named entity recognition [C]. In Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003-Volume 4. 2003: 142–147.

[45] Doddington G R, Mitchell A, Przybocki M A, et al. The Automatic Content Extraction (ACE) Program-Tasks, Data, and Evaluation. [C]. In LREC. 2004: 1.

[46] Rau L F. Extracting company names from text[C]//[1991] Proceedings. The Seventh IEEE Conference on Artificial Intelligence Application. IEEE, 1991, 1: 29-32.

[47] Settles B. Biomedical named entity recognition using conditional random fields and rich

AMiner

feature sets [C]. In Proceedings of the international joint workshop on natural language processing in biomedicine and its applications.2004:104–107.

[48] Zhou G, Su J. Named entity recognition using an HMM-based chunk tagger [C]. In proceedings of the 40th Annual Meeting on Association for Computational Linguistics. 2002: 473–480.

[49] Ittycheriah A, Lita L, Kambhatla N, et al. Identifying and tracking entity mentions in a maximum entropy framework [C]. In Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology: companion volume of the Proceedings of HLTNAACL2003–short papers-Volume 2. 2003: 40–42.

[50] Isozaki H, Kazawa H. Efficient support vector classifiers for named entity recognition [C]. In Proceedings of the 19th international conference on Computational linguistics-Volume1. 2002: 1–7.

[51] Guo H, Jiang J, Hu G, et al. Chinese named entity recognition based on multilevel linguistic features [C]. In International Conference on Natural Language Processing. 2004: 90–99.

[52] Zhou Y, Huang C, Gao J, et al. Transformation based Chinese entity detection and tracking [C]. In Companion Volume to the Proceedings of Conference including Posters/Demos and tutorial abstracts. 2005.

[53] Lu W, Roth D. Joint mention extraction and classification with mention hypergraphs [C]. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. 2015: 857–867.

[54] Florian R, Jing H, Kambhatla N, et al. Factorizing complex models: A case study in mention detection [C]. In Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics. 2006: 473–480.

[55] Li Q, Ji H. Incremental joint extraction of entity mentions and relations [C]. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2014: 402–412.

[56] Hacioglu K, Douglas B, Chen Y. Detection of entity mentions occurring in English and Chinese text [C]. In Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing. 2005: 379–386.

[57] Kambhatla N. Combining lexical, syntactic, and semantic features with maximum entropy models for extracting relations [C]. In Proceedings of the ACL 2004 on Interactive poster and demonstration sessions. 2004: 22.

[58] GuoDong Z, Jian S, Jie Z, et al. Exploring various knowledge in relation extraction [C]. In

AMiner

Proceedings of the 43rd annual meeting on association for computational linguistics. 2005: 427–434.
[59] Brin S. Extracting patterns and relations from the world wide web [C]. In International Workshop on The World Wide Web and Databases. 1998: 172–183.
[60] Sun A, Grishman R. Active learning for relation type extension with local and global data views [C]. In Proceedings of the 21st ACM international conference on Information and knowledge management. 2012: 1105–1112.
[61] Zhu X, Ghahramani Z. Learning from labeled and unlabeled data with label propagation [J]. 2002.
[62] Chen J, Ji D, Tan C L, et al. Relation extraction using label propagation based semi-supervised learning [C]. In Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics. 2006: 129–136.
[63] Banko M, Cafarella M J, Soderland S, et al. Open information extraction from the web. [C]. In IJCAI. 2007: 2670–2676.
[64] Banko M, Etzioni O. The tradeoffs between open and traditional relation extraction [J]. Proceedings of ACL-08: HLT. 2008: 28–36.
[65] Zhu J, Nie Z, Liu X, et al. StatSnowball: a statistical approach to extracting entity relationships [C]. In Proceedings of the 18th international conference on World wide web. 2009: 101–110.
[66] Wu F, Weld D S. Open information extraction using Wikipedia [C]. In Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics. 2010: 118–127.
[67] Fader A, Soderland S, Etzioni O. Identifying relations for open information extraction [C]. In Proceedings of the conference on empirical methods in natural language processing. 2011: 1535–1545.
[68] Gamallo P, Garcia M, Fernández-Lanza S. Dependency-based open information extraction [C]. In Proceedings of the Joint Workshop on Unsupervised and SemiSupervised Learning in NLP. 2012: 10–18.
[69] Surdeanu M, Ji H. Overview of the english slot filling track at the tac2014 knowledge base population evaluation [C]. In Proc. Text Analysis Conference (TAC2014). 2014.
[70] Pasca M, Van Durme B. What You Seek Is What You Get: Extraction of Class Attributes from Query Logs. [C]. In IJCAI. 2007: 2832–2837.
[71] Lee T, Wang Z, Wang H, et al. Attribute extraction and scoring: A probabilistic approach [C]. In Data Engineering (ICDE), 2013 IEEE 29th International Conference on. 2013: 194–205.
[72] Raju S, Pingali P, Varma V. An unsupervised approach to product attribute extraction [C]. In
AMiner

European Conference on Information Retrieval. 2009: 796–800.

[73] 沈晓卫, 李培峰, 朱巧明. 槽填充中抽取模式的优化方法 [J]. 中文信息学报.2015, 29 (2): 199–206.

[74] 斯坦福自然语言处理工具. https://nlp.stanford.edu/software/ lex-parser.html.

[75] Ye Z, Lin H, Su S, et al. Person attribute extracting based on svm [J]. Journal of Computer Research and Development. 2007, 44: 271–275.

[76] Huang R, Riloff E. Classifying message board posts with an extracted lexicon of patient attributes [C]. In Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing. 2013: 1557–1562.

[77] Liu B, Lane I. Attention-based recurrent neural network models for joint intent detection and slot filling [J]. arXiv preprint arXiv:1609.01454. 2016.

[78] Yao W, Liu J, Cai Z. Personal Attributes Extraction in Chinese Text Based on Distant-Supervision and LSTM [M] // Yao W, Liu J, Cai Z. Advances in Computer Science and Ubiquitous Computing. Springer, 2017: 2017: 511–515.

[79] Adel H, Roth B, Schütze H. Comparing convolutional neural networks to traditional models for slot filling [J]. arXiv preprint arXiv:1603.05157. 2016.

[80] 刘倩, 伍大勇, 刘悦等, 结合全局特征的命名实体属性值抽取 [J]. 计算机研究与发展. 2016, 53 (4): 941–948.

[81] Wu F, Weld D S. Autonomously semantifying wikipedia [C]. In Proceedings of the sixteenth ACM conference on Conference on information and knowledge management. 2007: 41–50.

[82] Probst K, Ghani R, Krema M, et al. Semi-Supervised Learning of Attribute-Value Pairs from Product Descriptions. [C]. In IJCAI. 2007: 2838–2843.

[83] Zhonghe H, Zhongcheng Z, Liang G, et al. Chinese entity attributes extraction based on Bidirectional LSTM Networks. [C]. In Computational Science and Engineering. 2016: 848–853.

[84] 曾道建, 来斯惟, 张元哲等. 面向非结构化文本的开放式实体属性抽取 [J]. 江西师范大学学报: 自然科学版. 2013, 37 (3): 279–283.

[85] Cucerzan S. Large-scale named entity disambiguation based on Wikipedia data [C]. In Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL). 2007.

[86] Ratinov L, Roth D, Downey D, et al. Local and global algorithms for disambiguation to wikipedia [C]. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1. 2011: 1375–1384.

[87] Bunescu R, Paşca M. Using encyclopedic knowledge for named entity disambiguation [C]. In 11th conference of the European Chapter of the Association for Computational Linguistics.

AMiner

2006.
[88] Han X, Zhao J. Named entity disambiguation by leveraging wikipedia semantic knowledge [C]. In Proceedings of the 18th ACM conference on Information and knowledge management. 2009: 215–224.
[89] Fang W, Zhang J, Wang D, et al. Entity disambiguation by knowledge and text jointly embedding [C]. In Proceedings of The 20th SIGNLL Conference on Computational Natural Language Learning. 2016: 260–269.
[90] Shen W, Wang J, Han J. Entity linking with a knowledge base: Issues, techniques, and solutions [J]. IEEE Transactions on Knowledge and Data Engineering. 2015, 27 (2): 443–460.
[91] 实体链接：从文本到概念. http://www.infoq.com/cn/presentations/entity-links.
[92] Han X, Sun L, Zhao J. Collective entity linking in web text: a graph-based method [C]. In Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval. 2011: 765–774.
[93] Gottipati S, Jiang J. Linking entities to a knowledge base with query expansion [C]. In Proceedings of the Conference on Empirical Methods in Natural Language Processing. 2011: 804–813.
[94] Shen W, Wang J, Luo P, et al. Linden: linking named entities with knowledge base via semantic knowledge [C]. In Proceedings of the 21st international conference on World Wide Web. 2012: 449–458.
[95] Zhang W, Sim Y C, Su J, et al. Entity linking with effective acronym expansion, instance selection, and topic modeling [C]. In IJCAI. 2011: 1909–1914.
[96] Dredze M, McNamee P, Rao D, et al. Entity disambiguation for knowledge base population [C]. In Proceedings of the 23rd International Conference on Computational Linguistics. 2010: 277–285.
[97] Monahan S, Lehmann J, Nyberg T, et al. Cross-Lingual Cross-Document Coreference with Entity Linking. [C]. In TAC. 2011.
[98] Lin T, Etzioni O, et al. Entity linking at web scale [C]. In Proceedings of the Joint WorkshoponAutomaticKnowledgeBaseConstructionandWeb-scaleKnowledge Extraction. 2012: 84–88.
[99] Hoffart J, Yosef M A, Bordino I, et al. Robust disambiguation of named entities in text [C]. In Proceedings of the Conference on Empirical Methods in Natural Language Processing. 2011: 782–792.
[100] Guo S, Chang M-W, Kiciman E. To link or not to link? a study on end-to-end tweet entity linking [C]. In Proceedings of the 2013 Conference of the North American Chapter of the
AMiner

Association for Computational Linguistics: Human Language Technologies. 2013: 1020–1030.
[101] Taylor Cassidy Z C, Artiles J, Ji H, et al. Cuny-uiuc-sri tac-kbp2011 entity linking system description [C]. In Proceedings of Text Analysis Conference. 2011.
[102] Liu X, Li Y, Wu H, et al. Entity linking for tweets [C]. In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2013: 1304–1311.
[103] Chen Z, Ji H. Collaborative ranking: A case study on entity linking [C]. In Proceedings of the Conference on Empirical Methods in Natural Language Processing. 2011: 771–781.
[104] Pilz A, Paaß G. From names to entities using thematic context distance [C]. In Proceedings of the 20th ACM international conference on Information and knowledge management. 2011: 857–866.
[105] Shen W, Wang J, Luo P, et al. LIEGE:: link entities in web lists with knowledge base [C]. In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. 2012: 1424–1432.
[106] Kulkarni S, Singh A, Ramakrishnan G, et al. Collective annotation of Wikipedia entities in web text [C]. In Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining. 2009: 457–466.
[107] Han X, Sun L. A generative entity-mention model for linking entities with knowledge base [C]. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1. 2011: 945– 954.
[108] Demartini G, Difallah D E, Cudré-Mauroux P. ZenCrowd: leveraging probabilistic reasoning and crowdsourcing techniques for large-scale entity linking [C]. In Proceedings of the 21st international conference on World Wide Web. 2012: 469– 478.
[109] Shen W, Wang J, Luo P, et al. Linking named entities in tweets with knowledge base via user interest modeling [C]. In Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. 2013: 68–76.
[110] Chen Z, Tamang S, Lee A, et al. CUNY-BLENDER TAC-KBP2010Entity Linking and Slot Filling System Description. [C]. In TAC. 2010.
[111] Varma V, Pingali P, Katragadda R, et al. IIIT Hyderabad at TAC 2009. [C]. In TAC. 2009.
[112] Jerry R. Hobbs. Resolving pronoun references[J]. Lingua, 44(4):0-338.
[113] Grosz B J, Weinstein S, Joshi A K. Centering: A framework for modeling the local coherence of discourse[J]. Computational linguistics, 1995, 21(2): 203-225.
[114] Mccarthy J F , Lehnert W G . Using Decision Trees for Coreference Resolution[J]. 1995.
[115] Bean D L, Riloff E. Unsupervised learning of contextual role knowledge for coreference

AMiner

resolution[C]//Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics: HLT-NAACL 2004. 2004: 297-304.
[116] Pantel P , Crestan E , Borkovsky A , et al. Web-Scale Distributional Similarity and Entity Set Expansion[C]// Conference on Empirical Methods in Natural Language Processing. DBLP, 2009.
[117] Wang C, Danilevsky M, Desai N, et al. A phrase mining framework for recursive construction of a topical hierarchy[C]// 2013.
[118] Song Y , Liu S , Wang H , et al. Automatic taxonomy construction from keywords:, 2016.
[119] Tan C H , Agichtein E , Ipeirotis P , et al. Trust, but Verify: Predicting Contribution Quality for Knowledge Base Construction and Curation[J]. 2014.
[120] Shi-Rui S , Jun-Hua X . Identifying Relations for Open Information Extraction.[C]// Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 2011.
[121] 王鑫，邹磊，王朝坤等，知识图谱数据管理研究综述[J]．软件学报, 2019(7).
[122] https://www.w3school.com.cn/rdf/index.asp.
[123] Etzioni O, Cafarella M, Downey D, et al. Methods for domain-independent information extraction from the web [C]. In Proceedings-Nineteenth National Conference on Artificial Intelligence (AAAI-2004): Sixteenth Innovative Applications of Artificial Intelligence Conference (IAAI-2004). 2004.
[124] https://www.w3.org/2001/sw/wiki/SPARQL.
[125] https://www.w3cschool.cn/neo4j/neo4j_cql_introduction.html.
[126] https://www.w3cschool.cn/neo4j/neo4j_features_advantages.html.
[127] Nilsson N J. Principles of artificial intelligence[M]. Morgan Kaufmann, 2014.
[128] Kompridis N. So We Need Something Else for Reason to Mean[J]. International Journal of Philosophical Studies, 2000, 8(3): 271-295.
[129] Luis Tari. Knowledge Inference[J]. 2013.
[130] Chen X , Jia S , Xiang Y . A review: Knowledge reasoning over knowledge graph[J]. Expert Systems with Applications, 2019, 141:112948.
[131] 杜政晓．基于认知计算的深度学习推理初探[D]．清华大学, 2020.
[132] CSDN.知识图谱入门（七）知识推理[EB/OL]. https://blog.csdn.net/pelhans/article/details/80091322
[133] Yang F , Yang Z , Cohen W W . Differentiable Learning of Logical Rules for Knowledge Base Reasoning[J]. 2017.
[134] 官赛萍，靳小龙，贾岩涛,等．面向知识图谱的知识推理研究进展[J]．软件学报, 2018,

AMiner

29(10):74-102.

[135] Wang W Y , Mazaitis K , Lao N , et al. Efficient Inference and Learning in a Large Knowledge Base: Reasoning with Extracted Information using a Locally Groundable First-Order Probabilistic Logic[J]. Computer ence, 2014.

[136] Cohen W W . TensorLog: A Differentiable Deductive Database[J]. 2016.

[137] Lao N , Cohen W W . Relational Retrieval Using a Combination of Path-Constrained Random Walks[J]. Machine Learning, 2010, 81(1):53-67.

[138] Luis Antonio Galárraga, Teflioudi C , Hose K , et al. AMIE: Association rule mining under incomplete evidence in ontological knowledge bases[C]// International Conference on World Wide Web. ACM, 2013.

[139] 官赛萍, 靳小龙, 贾岩涛等, 面向知识图谱的知识推理研究进展[J]. 软件学报, 2018, 29(10):74-102.

[140] Lin Y, Liu Z, Sun M, et al. Learning entity and relation embeddings for knowledge graph completion[C]. national conference on artificial intelligence, 2015: 2181-2187.

[141] Xiao H, Huang M, Hao Y, et al. TransG : A Generative Mixture Model for Knowledge Graph Embedding[J]. arXiv: Computation and Language, 2015.

[142] Ruobing Xie, Zhiyuan Liu, Maosong Sun. Representation learning of knowledge graphs with hierarchical types[C]// International Joint Conference on Artificial Intelligence. AAAI Press, 2016.

[143] Nguyen D Q, Sirts K, Qu L, et al. Neighborhood Mixture Model for Knowledge Base Completion[C]. conference on computational natural language learning, 2016: 40-50.

[144] Zhigang Wang, Juanzi Li. Text-enhanced representation learning for knowledge graph[C]// International Joint Conference on Artificial Intelligence. AAAI Press, 2016.

[145] Jiang T, Liu T, Ge T, et al. Towards Time-Aware Knowledge Graph Completion[C]. international conference on computational linguistics, 2016: 1715-1724.

[146] Yi Tay, Anh Tuan Luu, Siu Cheung Hui,等. Random Semantic Tensor Ensemble for Scalable Knowledge Graph Link Prediction[J]. 2017.

[147] Chang K W , Yih W T , Yang B , et al. Typed Tensor Decomposition of Knowledge Bases for Relation Extraction[C]// Emnlp. 2014.

[148] Nickel M , Rosasco L , Poggio T . Holographic Embeddings of Knowledge Graphs[J]. 2015.

[149] Qiao L, Minghao H, Xiaohui Y, et al. Representation Learning Based Relational Inference Algorithm with Semantical Aspect Awareness[J]. Journal of Computer Research and Development, 2017, 54(8).

[150] Liu H, Wu Y, Yang Y, et al. Analogical Inference for Multi-relational Embeddings[C]. international conference on machine learning, 2017: 2168-2178.

AMiner

[151] Socher R, Chen D, Manning C D, et al. Reasoning with neural tensor networks for knowledge base completion[C]//Advances in neural information processing systems. 2013: 926-934.
[152] Shi B , Weninger T . ProjE: Embedding Projection for Knowledge Graph Completion[J]. 2016.
[153] Xie R, Liu Z, Jia J, et al. Representation learning of knowledge graphs with entity descriptions[C]. national conference on artificial intelligence, 2016: 2659-2665.
[154] Neelakantan A, Roth B, Mccallum A, et al. Compositional Vector Space Models for Knowledge Base Inference.[C]. national conference on artificial intelligence, 2015.
[155] Guo L , Zhang Q , Ge W , et al. DSKG: A Deep Sequential Model for Knowledge Graph Completion[J]. 2018.
[156] Das R, Neelakantan A, Belanger D, et al. Chains of reasoning over entities, relations, and text using recurrent neural networks[J]. arXiv preprint arXiv:1607.01426, 2016.
[157] Xiong W , Hoang T , Wang W Y . DeepPath: A Reinforcement Learning Method for Knowledge Graph Reasoning[J]. 2017.
[158] Das R , Dhuliawala S , Zaheer M , et al. Go for a Walk and Arrive at the Answer: Reasoning Over Paths in Knowledge Bases using Reinforcement Learning[J]. 2017.
[159] 产品经理的人工智能学习库.自然语言生成 – Natural-language generation | NLG[EB/OL]. https://easyai.tech/ai-definition/nlg/
[160] 张建华, 陈家骏. 自然语言生成综述[J]. 计算机应用研究, 2006(08):7-9+19.
[161] 王昊奋, 邵浩. 自然语言处理实践：聊天机器人技术原理与应用[M]. 电子工业出版社. 2019.
[162] Peters M E , Neumann M , Iyyer M , et al. Deep contextualized word representations[J]. 2018.
[163] 知乎.【NLP】Transformer 详解[EB/OL].https://zhuanlan.zhihu.com/p/44121378
[164] Radford, Alec, et al. Improving language understanding by generative pre-training[J]. 2018(12).
[165] Devlin, Jacob, et al. Bert: Pre-training of deep bidirectional transformers for language understanding[J]. arXiv preprint. 2018.
[166] Yang Z, Dai Z, Yang Y, et al. Xlnet: Generalized autoregressive pretraining for language understanding[C]//Advances in neural information processing systems. 2019: 5753-5763.
[167] 博客园.XLNet 原理探究[EB/OL].

AMiner

https://www.cnblogs.com/huangyc/p/11446575.html#_label2

[168] 博客园.NLP 自然语言处理之综述[EB/OL]. https://www.cnblogs.com/qccz123456/p/11623990.html.

[169] 维基百科. 自然语言生成[EB/OL]. https://zh.wikipedia.org/wiki/%E8%87%AA%E7%84%B6%E8%AF%AD%E8%A8%80%E7%94%9F%E6%88%90.

[170] 袁江林, 郭志刚, 陈刚,等. 基于深度学习的文本自动生成技术研究综述[J]. 信息工程大学学报, 2018(5).

[171] 人工智能. 自然语言处理的技术难点与挑战_发展现状分析[EB/OL]. http://m.elecfans.com/article/608966.html.

[172] 维基百科. Latent Dirichlet Allocation [EB/OL]. https://en.wikipedia.org/wiki/Latent_Dirichlet_allocation

[173] Singhal, Amit. Introducing the knowledge graph: things, not strings[EB/OL]. Official google blog 5. 2012.

[174] Grishman, Ralph. Information extraction: Techniques and challenges[J]. International summer school on information extraction. 1997.

[175] Etzioni, Oren, et al. Open information extraction from the web[J]. Communications of the ACM. 2008: 68-74

[176] Nadeau, David, and Satoshi Sekine. A survey of named entity recognition and classification[J]. Lingvisticae Investigationes. 2007: 3-26.

[177] Lin, Yankai, et al. Learning entity and relation embeddings for knowledge graph completion[J]. Twenty-ninth AAAI conference on artificial intelligence. 2015.

[178] Bordes, Antoine, et al. Translating embeddings for modeling multi-relational data[J]. Advances in neural information processing systems. 2013.

[179] Reiter, Ehud, and Robert Dale. Building natural language generation systems[J]. Cambridge university press, 2000.

[180] Dong, Li, et al. Unified language model pre-training for natural language understanding and generation[J]. Advances in Neural Information Processing Systems. 2019.

[181] Vaswani, Ashish, et al. Attention is all you need[J]. Advances in neural information

AMiner

processing systems. 2017.

[182] CSDN. 肖仰华：知识图谱与认知智能[EB/OL]. https://blog.csdn.net/omnispace/article/details/80268053

[183] CSDN. 人机交互与智能的再思考[EB/OL]. https://blog.csdn.net/VucNdnrzk8iwX/article/details/90586278

[184] 张钹. 人工智能进入后深度学习时代[J]. 智能科学与技术学报, 2019.

AMiner

AMiner

附录 1 认知图谱相关的关键词列表

表 14 认知图谱相关领域的关键词列表

领域关键词	领域关键词	领域关键词
world wide web	random walk inference	syntactic generation
semantic network	probability reasoning	sentence summarization
resource description framework	ontology reasoning	report generator
ontology	neural logic network	product description generation
cognitive graph	neural logic inductive learning	neural response generation
knowledge graph	neural conversation model	neural machine translation
knowledge base	logical inference	neural language generation
graph representation learning	logic neural network	neural conversational model
yago	knowledge reasoning	neural conversation model
wikidata	knowledge graph reasoning	natural language generation
semantic web	first-order predicate reasoning	machine translation
opencyc	explainable reasoning	language model
graph level representation	entity inference	Generative Pretrained Transformer
graph kernels	description logic reasoning	generative text model
freebase	deep temporal reasoning	generating text with logic
dbpedia	deep reasoning network	discourse structure
variational reasoning	deep logic reasoning	dialogue systems
symbolic reasoning	deep logic network	dialogue response generation
relational neural network	commonsense reasoning	deep generative model
relation inference	action reasoning	abstractive summarization
reasoning with neural tensor network	text generation	text realization

AMiner

附录 2 代表性期刊和会议列表

表 15 认知图谱领域代表性期刊和会议列表

序号	期刊/会议全称	缩写
1	ACM Transactions on Algorithms	TALG
2	ACM Transactions on Computational Logic	TOCL
3	ACM Transactions on Computer-Human Interaction	TOCHI
4	ACM Transactions on Database Systems	TODS
5	ACM Transactions on Graphics	TOG
6	ACM Transactions on Information Systems	TOIS
7	ACM Transactions on Knowledge Discovery from Data	TKDD
8	ACM Transactions on Mathematical Software	TOMS
9	ACM Transactions on Programming Languages & Systems	TOPLAS
10	ACM Transactions on Software Engineering and Methodology	TOSEM
11	ACM Transactions on the Web	TWEB
12	Advanced Engineering Informatics	AEI
13	Algorithmica	Algorithmica
14	Applied Intelligence	
15	Artificial Intelligence	AI
16	Artificial Intelligence in Medicine	AIM
17	Automated Software Engineering	ASE
18	Autonomous Agents and Multi-Agent Systems	AAMAS
19	Behaviour & Information Technology	BIT
20	Bioinformatics	
21	BMC Bioinformatics	
22	Briefings in Bioinformatics	
23	Computational complexity	CC
24	Computational Linguistics	
25	Computer Aided Geometric Design	CAGD
26	Computer Graphics Forum	CGF
27	Computer Speech and Language	
28	Computer Supported Cooperative Work	CSCW
29	Computer Vision and Image Understanding	CVIU
30	Computer-Aided Design	CAD
31	Computers & Graphics	CG
32	Data and Knowledge Engineering	DKE
33	Data Mining and Knowledge Discovery	DMKD
34	Empirical Software Engineering	ESE
35	Engineering Applications of Artificial Intelligence	EAAI

AMiner

36	European Journal of Information Systems	EJIS
37	Evolutionary Computation	
38	Expert Systems with Applications	Expert Syst. Appl.
39	Formal Aspects of Computing	FAC
40	Formal Methods in System Design	FMSD
41	Future Generation Computer Systems	FGCS
42	GeoInformatica	
43	Human Computer Interaction	HCI
44	IEEE Geoscience and Remote Sensing Letters	GRSL
45	IEEE Journal of Biomedical and Health Informatics	JBHI
46	IEEE Signal Processing Letters	SPL
47	IEEE Trans on Pattern Analysis and Machine Intelligence	TPAMI
48	IEEE Transactions on Affective Computing	TAC
49	IEEE Transactions on Audio Speech, and Language Processing	TASLP
50	IEEE Transactions on Automation Science and Engineering	TASAE
51	IEEE Transactions on Big Data	TBD
52	IEEE Transactions on Circuits and Systems for Video Technology	TCSVT
53	IEEE Transactions on Cybernetics	
54	IEEE Transactions on Evolutionary Computation	TEC
55	IEEE Transactions on Fuzzy Systems	TFS
56	IEEE Transactions on Games	TG
57	IEEE Transactions on Geoscience and Remote Sensing	TGARS
58	IEEE Transactions on Human-Machine Systems	
59	IEEE Transactions on Image Processing	TIP
60	IEEE Transactions on Information Theory	TIT
61	IEEE Transactions on Intelligent Transportation Systems	TITS
62	IEEE Transactions on Knowledge and Data Engineering	TKDE
63	IEEE Transactions on Medical Imaging	TMI
64	IEEE Transactions on Multimedia	TMM
65	IEEE Transactions on Neural Networks and learning systems	TNNLS
66	IEEE Transactions on Robotics	TR
67	IEEE Transactions on Service Computing	TSC
68	IEEE Transactions on Software Engineering	TSE
69	IEEE Transactions on Visualization and Computer Graphics	TVCG
70	IEEE-ACM Transactions on Computational Biology and Bioinformatics	TCBB
71	Image and Vision Computing	
72	Information and Computation	IANDC
73	Information and Management	I&M
74	Information and Software Technology	IST
75	Information Processing and Management	IPM

76	Information Sciences	
77	Information Systems	IS
78	INFORMS Journal on Computing	INFORMS
79	Interacting with Computers	IWC
80	International Journal of Approximate Reasoning	IJAR
81	International Journal of Computer Vision	IJCV
82	International Journal of Geographical Information Science	IJGIS
83	International Journal of Human Computer Studies	IJHCS
84	International Journal of Human-Computer Interaction	IJHCI
85	International Journal of Intelligent Systems	IJIS
86	International Journal of Intelligent Systems	IJIS
87	International Journal of Neural Systems	
88	International Journal on Software Tools for Technology Transfer	STTT
89	Journal of Artificial Intelligence Research	JAIR
90	Journal of Automated Reasoning	
91	Journal of Biomedical Informatics	JBI
92	Journal of Computer and System Sciences	JCSS
93	Journal of Computer Science and Technology	JCST
94	Journal of Global Optimization	JGO
95	Journal of Grid Computing	JGC
96	Journal of Machine Learning Research	JMLR
97	Journal of Speech, Language, and Hearing Research	JSLHR
98	Journal of Strategic Information Systems	J. Strategic Inf. Sys.
99	Journal of Symbolic Computation	JSC
100	Journal of Systems and Software	JSS
101	Journal of the ACM	JACM
102	Journal of the American Medical Informatics Association	JAMIA
103	Journal of the American Society for Information Science and Technology	JASIST
104	Journal of Web Semantics	JWS
105	Knowledge and Information Systems	KAIS
106	KNOWLEDGE-BASED SYSTEMS	
107	Machine Learning	
108	Neural Computation	
109	Neural Networks	
110	Neurocomputing	
111	Pattern Recognition	PR
112	Pattern Recognition Letters	PRL
113	Pervasive and Mobile Computing	PMC
114	Proceedings of the IEEE	Proc. IEEE
115	Requirements Engineering	RE

AMiner

116	Science China	
117	Science of Computer Programming	SCP
118	SIAM Journal on Computing	SICOMP
119	SIAM Journal on Imaging Sciences	SIIMS
120	Signal Processing	
121	Software and System Modeling	SoSyM
122	Software Testing, Verification and Reliability	STVR
123	Software: Practice and Experience	SPE
124	Speech Communication	Speech Com
125	The VLDB Journal	VLDBJ
126	Theoretical Computer Science	TCS
127	Transactions of the Association for Computational Linguistics	TACL
128	Tsinghua Science and Technology	
129	User Modeling and User-Adapted Interaction	UMUAI
130	电子学报	
131	计算机辅助设计与图形学学报	
132	计算机学报	
133	计算机研究与发展	
134	清华大学学报（自然科学版）	
135	软件学报	
136	图学学报	
137	系统仿真学报	
138	智能系统学报	
139	中国科学	
140	中国图像图形学报	
141	中文信息学报	
142	自动化学报	
143	AAAI Conference on Artificial Intelligence	AAAI
144	ACM Conference on Computer Supported Cooperative Work and Social Computing	CSCW
145	ACM Conference on Designing Interactive Systems	DIS
146	ACM Conference on Economics and Computation	EC
147	ACM Conference on Human Factors in Computing Systems	CHI
148	ACM Conference on Management of Data	SIGMOD
149	ACM International Conference on Information and Knowledge Management	CIKM
150	ACM International Conference on Intelligent User Interfaces	IUI
151	ACM International Conference on Multimedia	ACM MM
152	ACM International Conference on Multimodal Interaction	ICMI
153	ACM International Conference on Ubiquitous Computing	UbiComp
154	ACM International Conference on Web Search and Data Mining	WSDM
155	ACM Knowledge Discovery and Data Mining	SIGKDD

AMiner

156	ACM SIGGRAPH Annual Conference	SIGGRAPH
157	ACM SIGMM International Conference on Multimedia Retrieval	ICMR
158	ACM SIGPLAN Conference on Programming Language Design & Implementation	PLDI
159	ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages	POPL
160	ACM SIGSOFT Symposium on the Foundation of Software Engineering/ European Software Engineering Conference	FSE/ESEC
161	ACM Symposium on Computational Geometry	SoCG
162	ACM Symposium on Interactive 3D Graphics	SI3D
163	ACM Symposium on Operating Systems Principles	SOSP
164	ACM Symposium on Principles of Database Systems	PODS
165	ACM Symposium on the Theory of Computing	STOC
166	ACM Symposium on User Interface Software and Technology	UIST
167	ACM/Eurographics Symposium on Computer Animation	SCA
168	ACM-SIAM Symposium on Discrete Algorithms	SODA
169	Annual Conference on Computational Learning Theory	COLT
170	Annual Conference on Neural Information Processing Systems	NeurIPS
171	Annual Meeting of the Association for Computational Linguistics	ACL
172	Artificial Intelligence and Statistics	AISTATS
173	British Machine Vision Conference	BMVC
174	Computer Aided Verification	CAV
175	Computer Science Logic	CSL
176	Conference on Empirical Methods in Natural Language Processing	EMNLP
177	Conference on Object-Oriented Programming Systems, Languages, and Applications	OOPSLA
178	Conference on Recommender Systems	RecSys
179	Data Compression Conference	DCC
180	Database Systems for Advanced Applications	DASFAA
181	Eurographics	EG
182	Eurographics Conference on Visualization	EuroVis
183	Eurographics Symposium on Geometry Processing	SGP
184	Eurographics Symposium on Rendering	EGSR
185	European Conference on Computer Vision	ECCV
186	European Conference on IR Research	ECIR
187	European Conference on Object-Oriented Programming	ECOOP
188	European Joint Conferences on Theory and Practice of Software	ETAPS
189	European Symposium on Algorithms	ESA
190	Evaluation and Assessment in Software Engineering	EASE
191	Formal Methods in Computer-Aided Design	FMCAD
192	Genetic and Evolutionary Computation Conference	GECCO
193	IEEE Annual Symposium on Foundations of Computer Science	FOCS

AMiner

194	IEEE Conference on Computational Complexity	CCC
195	IEEE Conference on Computer Vision and Pattern Recognition	CVPR
196	IEEE International Conference on Acoustics, Speech and SP	ICASSP
197	IEEE International Conference on Data Engineering	ICDE
198	IEEE International Conference on Pervasive Computing and Communications	PERCOM
199	IEEE International Conference on Program Comprehension	ICPC
200	IEEE International Conference on Robotics and Automation	ICRA
201	IEEE International Conference on Software Testing, Verification and Validation	ICST
202	IEEE International Requirement Engineering Conference	RE
203	IEEE International Semantic Web Conference	ISWC
204	IEEE International Symposium on Information Theory	ISIT
205	IEEE International Symposium on Performance Analysis of Systems and Software	ISPASS
206	IEEE Pacific Visualization Symposium	PacificVis
207	IEEE Symposium on Logic in Computer Science	LICS
208	IEEE Virtual Reality	VR
209	IEEE Visualization Conference	IEEE VIS
210	IEEE\RSJ International Conference on Intelligent Robots and Systems	IROS
211	Innovations in Theoretical Computer Science	ITCS
212	International Colloquium on Automata, Languages and Programming	ICALP
213	International Conference on Advanced Information Systems Engineering	CAiSE
214	International Conference on Algorithmic Learning Theory	ALT
215	International Conference on Automated Deduction/International Joint Conference on Automated Reasoning	CADE/IJCAR
216	International Conference on Automated Planning and Scheduling	ICAPS
217	International Conference on Automated Software Engineering	ASE
218	International Conference on Automatic Face and Gesture Recognition	FG
219	International Conference on Business Process Management	BPM
220	International Conference on Computational Linguistics	COLING
221	International Conference on Computer Vision	ICCV
222	International Conference on Concurrency Theory	CONCUR
223	International Conference on Data Mining	ICDM
224	International Conference on Database Theory	ICDT
225	International Conference on Document Analysis and Recognition	ICDAR
226	International Conference on Embedded Software	EMSOFT
227	International Conference on Extending Database Technology	EDBT
228	International Conference on Function Programming	ICFP

AMiner

229	International Conference on Human Computer Interaction with Mobile Devices and Services	MobileHCI
230	International Conference on Hybrid Systems: Computation and Control	HSCC
231	International Conference on Image Processing	ICIP
232	International Conference on Innovative Data Systems Research	CIDR
233	International conference on Intelligent Systems for Molecular Biology	ISMB
234	International Conference on Learning Representations	ICLR
235	International Conference on Machine Learning	ICML
236	International Conference on Medical Image Computing and Computer Assisted Intervention	MICCAI
237	International Conference on Model Driven Engineering Languages and Systems	MoDELS
238	International Conference on Pattern Recognition	ICPR
239	International Conference on Principles and Practice of Constraint Programming	CP
240	International Conference on Research in Computational Molecular Biology	RECOMB
241	International Conference on Research on Development in Information Retrieval	SIGIR
242	International Conference on Service Computing	SCC
243	International Conference on Software Analysis, Evolution, and Reengineering	SANER
244	International Conference on Software Engineering	ICSE
245	International Conference on Software Maintenance and Evolution	ICSME
246	International Conference on Uncertainty in Artificial Intelligence	UAI
247	International Conference on Verification, Model Checking, and Abstract Interpretation	VMCAI
248	International Conference on Very Large Data Bases	VLDB
249	International Conference on Web Services （Research Track）	ICWS
250	International Joint Conference on Artificial Intelligence	IJCAI
251	International Joint Conference on Autonomous Agents and Multi-agent Systems	AAMAS
252	International Joint Conference on Biometrics	ICB
253	International Joint Conference on Neural Networks	IJCNN
254	International Middleware Conference	Middleware
255	International Symposium on Automated Technology for Verification and Analysis	ATVA
256	International Symposium on Empirical Software Engineering and Measurement	ESEM

AMiner

257	International Symposium on Formal Methods	FM
258	International Symposium on Mixed and Augmented Reality	ISMAR
259	International Symposium on Software Reliability Engineering	ISSRE
260	International Symposium on Software Testing and Analysis	ISSTA
261	International World Wide Web Conferences	WWW
262	Mining Software Repositories	MSR
263	Pacific Graphics: The Pacific Conference on Computer Graphics and Applications	PG
264	Robotics: Science and Systems ·A Robotics Conference	RSS
265	Scandinavian Symposium and Workshops on Algorithm Theory/ the Algorithms and Data Structures Symposium	SWAT/WADS
266	SIAM International Conference on Data Mining	SDM
267	Symposium on Solid and Physical Modeling	SPM
268	Symposium on Theoretical Aspects of Computer Science	STACS
269	The Annual Conference of the North American Chapter of the Association for Computational Linguistics	NAACL

AMiner

附录 3 国家自然科学基金 NSFC 项目

表 16 认知图谱领域国家自然科学基金支持的相关项目（2010-2020 年）

项目类别	项目标题	依托单位
专项基金项目	Domain 理论在不确定性推理中的应用	电子科技大学
	混合贝叶斯网的概率推理	东北师范大学
	线性时序关系下推理的概率计量化模型	陕西师范大学
	逻辑等价算子在不确定性推理中的应用	宝鸡文理学院
	基于本体的地理信息语义转换理论与方法研究	中国人民解放军信息工程大学
	面向移动计算的服务不确定性模型及其推理技术	华南师范大学
	面向机器理解的汉语隐喻计算研究与知识库建设	北京语言大学
	数据密集型计算与非结构化数据管理暑期学校	中国人民大学
	自然科学基金管理中评议人自动指派方法研究	香港城市大学深圳研究院
	基础科研知识库建设关键技术研究	浙江大学
重点项目	复杂推理的序、代数和逻辑方法及其计量化模型	陕西师范大学
	基于本体的多策略民汉机器翻译研究	北京理工大学
	大数据驱动的智慧医疗健康管理创新	北京交通大学
	面向资源稀缺型语言的机器翻译理论方法及关键技术研究	昆明理工大学
	面向领域大数据的事件知识图谱构建研究	苏州大学
	面向领域大数据的知识图谱构建	中国科学院计算机网络信息中心
重大研究计划	面向文本推理的汉语语义计算模型研究	北京大学
	非常规突发事件应急管理本体建模与时空数据集成研究	中国人民大学
	黑河流域水资源综合管理决策支持系统集成研究	中国科学院地理科学与资源研究所
	基于人类驾驶知识的无人驾驶车辆智能决策系统研究	中国科学院合肥物质科学研究院
	城区真实交通环境无人驾驶车辆关键技术与平台研究	北京理工大学
	基于高效电子、质子和能量迁移的功能组装体	中国科学院理化技术研究所
	面向汉语文本理解的语义计算方法	哈尔滨工业大学
	距离敏感的电调控化学刻蚀纳米加工新方法的研究	厦门大学
	非结构化管理决策大数据平台构建与关键技术	北京工业大学
	支持应急决策的气象灾害大数据融合的方法研究	南京信息工程大学

AMiner

	商用客机气动噪声大规模并行计算的建模、算法与软件	上海大学
	融合多源社会化网络数据的个体价值发现与预测	山东大学
	基于深度学习的多源异构质检大数据融合与分析技术研究	中国标准化研究院
	面向全流程智慧健康管理决策的多源异构大数据融合方法研究	清华大学
	大数据驱动的全景式个性化心血管健康管理研究	上海交通大学
	基于多元智能材料的超轻质软体机械臂大变形致动及变刚度机理研究	哈尔滨工业大学
	蒙皮镜像加工机器人系统的主动顺应与协调控制研究	上海交通大学
	面向商务大数据的知识图谱引擎构建方法与关键技术研究	浙江大学
优秀青年科学基金项目	绿色交通需求管理政策研究	北京交通大学
应急管理项目	大数据环境下的管理决策理论动态研究	同济大学
	大数据时代的生物信息学发展战略研究	中国科学院生物物理研究所
	中美信息科学资助格局比较研究——以中国 NSFC 和美国 NSF 为例	北京大学
	创新群体团队合作机制研究	哈尔滨工业大学
	工程科技发展战略研究知识体系与支持系统	中国航天系统科学与工程研究院
	开放科学驱动下的开放获取政策研究与平台构建	北京大学
	基金项目拆分融合相似性检测关键技术	中南大学
青年科学基金项目	描述逻辑及其扩展中的相容性问题和推理算法研究	湖南大学
	不确定性推理与语义网中知识表示的数学基础	电子科技大学
	基于动态时序语义的逻辑推理及其量化模型	陕西师范大学
	语篇话题结构对语义整合的影响及其认知神经机制	中国科学院心理研究所
	基于粗糙集理论的地理信息服务演绎推理研究	南京邮电大学
	基于认知实验的地图汉字注记配置规则及其形式化表达研究	西北农林科技大学
	基于语义网络的街区场景相似性研究	长安大学
	全球先验 BRDF 知识库的构建及高分辨率地表反照率的反演	天津师范大学
	基于视觉信息感知与操纵知识推理的桥区船舶航行安全研究	武汉理工大学
	基于设计认知的创意知识表达与获取方法研究	浙江大学
	TRIZ 多冲突设计问题的工程语义网求解及专利知识挖掘方法与应用	大连理工大学

AMiner

	基于认知语言学原理的工艺知识发现方法研究	西北工业大学
	基于云知识库的智能数控系统加工工艺规划方法研究	山东大学
	时空耦合动态特征下铁路站场 BIM 超图语义本体建模与一致性推理方法	中南大学
	基于本体的视频语义内容分析方法研究	中国人民解放军国防科学技术大学
	结合描述逻辑和霍恩规则的不确定推理算法	复旦大学
	基于描述逻辑和模型检测的行动理论研究	桂林电子科技大学
	基于主动学习的半监督领域本体自动构建	吉林大学
	基于领域本体的 Petri 网自动集成机理与应用模式研究	济南大学
	基于语义的稳定本体测量和评估研究	华北电力大学
	基于 Markov 逻辑网络的限定领域中文自动问答系统研究	复旦大学
	面向网络舆论的动态本体学习模型研究	清华大学
	面向语义 Web 复杂应用环境的本体映射关键技术研究	东南大学
	基于描述逻辑的本体调试关键技术的研究	东南大学
	面向领域本体的深度学习方法研究	中国人民大学
	基于网络预警机模型的认知网络 QoS 自主性管理与控制关键技术研究	南京邮电大学
	修复不一致 OWL 本体的高效方法研究	广东外语外贸大学
	不确定性推理的 Borel 型概率计量化模型	陕西师范大学
	软约束满足问题的符号表示及其推理研究	桂林电子科技大学
	Web 规模 RDF 图数据的高效率路径查询及推理研究	天津大学
	基于用户兴趣的 Web 知识检索与推理双向融合研究	中国科学院自动化研究所
	面向自然语言文本生成的事件语义计算研究	武汉科技大学
	基于本体的物联网信息智能管理机制研究	湖南农业大学
	基于本体的交互协作电视跨媒体语义融合研究	上海交通大学
	领域本体服务组合的 Petri 网建模和非功能特性分析	重庆大学
	引入涉身认知机制的汉语隐喻计算模型及其实现	杭州电子科技大学
	粗糙逻辑的程度化与随机化推理模型	西安石油大学
	藏语语义本体的概念识别和上下位关系获取技术研究	中央民族大学
	本体的保守扩充及其推理机制研究	广西师范大学
	基于 PAM 概率主题模型的本体概念及层次结构学习研究	同济大学
	基于直觉模糊命题逻辑的语言真值归结自动推理研究	辽宁师范大学

AMiner

	海量、动态、嘈杂语义数据集上的递增随时推理方法研究	西北工业大学
	基于 Wiki 资源的中英文跨语言本体知识库构建	北京师范大学
	基于数据驱动的中文自然语言生成关键技术研究	北京邮电大学
	基于本体的认知推理模型及其应用研究	电子科技大学
	基于模糊数据库的大规模复杂模糊本体自动构建与存储关键技术研究	东北大学
	基于自然语言处理技术的生物实体语义网络研究和应用	华中农业大学
	基于公理模糊集与粒化认知机理的形式概念分析研究	大连海事大学
	命题与模态逻辑的扩展规则推理与混合推理方法研究	吉林大学
	基于生物医学文献和领域本体的蛋白质复合物预测方法研究	大连理工大学
	基于高阶逻辑的归纳逻辑程序设计学习算法及其应用研究	东北林业大学
	基于弱指导机器学习技术的中文领域本体非分类关系自动学习研究	河北科技大学
	模糊本体知识建模方法与语义信息服务策略研究	中南大学
	可计算性逻辑中若干 Cirquent 演算系统的研究	西安电子科技大学
	基于搜索过程知识表示与推理的进化多目标优化算法研究	西安电子科技大学
	面向大数据深度分析的马尔科夫逻辑理论与算法研究	中国科学院自动化研究所
	基于描述逻辑的事件推理关键问题研究	上海大学
	基于动态描述逻辑的异构数据库数据整合技术研究	中国科学院软件研究所
	对等计算环境中多维认知与动态推理的分布式信任机制研究	东北大学
	需求跟踪行为认知与主动学习机理研究	福建工程学院
	语义 Web 模糊规则互换与推理关键技术研究	辽宁工程技术大学
	面向资源型社交网站的知识图谱构建方法研究	南华大学
	基于变分推理的马尔可夫随机场可近似性层次结构研究	天津科技大学
	基于矩阵—张量协同分解的大规模知识推理方法研究	中国科学院信息工程研究所
	面向专利大数据的机器学习与推理研究	中国科学院自动化研究所
	基于逻辑进程范型的空间随机仿真并行化研究	中国人民解放军63928 部队
	基于认知循环的任务关键系统可生存性自主增长模型与方法	哈尔滨理工大学

AMiner

	基于描述逻辑的复杂本体非经典知识表示和推理研究	南京大学
	基于知识图谱的服务机器人任务自动生成方法的研究与实现	中国科学技术大学
	面向事件预测的因果知识发现、验证和推理研究	中国科学院信息工程研究所
	知识图谱推理的理论与方法研究	北京航空航天大学
	结合知识图谱的概率话题模型研究	重庆邮电大学
	融合事件关系推理和情感博弈的网络不实信息演化机理研究	菏泽学院
	OWL 本体的协同调试问题研究	吉林大学
	大型复杂医学领域本体质量评估理论研究	南华大学
	面向知识库的实体链接技术研究	南开大学
	大规模 WiFi 轨迹隐含知识图谱挖掘研究	清华大学
	面向大规模网络评论文本的产品知识库构建关键技术研究	中国科学院自动化研究所
	基于开源生态知识图谱的软件舆情分析研究	中国人民解放军国防科学技术大学
	基于进化算法的大规模本体匹配问题研究	福建工程学院
	基于知识库构建的图像和视频角色语义关系的研究	山东大学
	基于深度学习和马尔科夫逻辑网络的特殊视频识别研究	中央民族大学
	物联网搜索中异构本体的语义融合研究	北京信息科技大学
	医学诊治过程中知识推理的关键技术研究	北京大学深圳研究生院
	基于图模式的知识图谱一致性问题研究	北京航空航天大学
	基于语义数据的迭代本体学习方法研究	南京邮电大学
	基于生成模型的本体学习方法研究	南京邮电大学
	基于外存的海量知识图谱数据的查询处理	香港浸会大学深圳研究院
	基于记忆、推理和注意力机制的端到端神经对话系统研究	中国科学院自动化研究所
	基于公共安全知识图谱的异常事件检测研究	中国科学院自动化研究所
	多智能体模态逻辑中的遗忘理论及其应用	暨南大学
	基于超网络模型的科学知识图谱分析研究	上海理工大学
	WEB2.0 环境下基于本体学习的观点挖掘研究	中南财经政法大学
	基于语义网格的数字图书馆个性化推荐模型研究	湖北工业大学

AMiner

	基于本体与异质复杂网络的推荐系统研究	北京科技大学
	基于通用加权 XML 模型的个性化用户兴趣本体研究	南京财经大学
	中文领域本体学习及半自动构建方法研究	福州大学
	基于可信语义 Wiki 的知识库构建方法与应用研究	西安电子科技大学
	基于 Web 的农产品安全信息传播与干预模型研究	北京市农林科学院
	基于多案例库协同推理的老年人健康评估方法研究	合肥工业大学
	基于影响图的 IT 项目风险分析及决策模型	河北科技大学
	建筑工程质量管理规范建模与检索的本体化方法研究	华中科技大学
	基于认知程度的双论域粗糙集理论及在动态决策中的应用研究	上海工程技术大学
	众包环境下基于胜任力的供需双方匹配研究	上海交通大学
	面向集成与交互的产品服务系统方案配置设计与评价方法研究	上海理工大学
	物流配送云资源虚拟化管理与服务组合优化研究	重庆大学
	电力行业有限需求信息下的间断需求预测研究	上海大学
	网站质量影响消费者在线购买决策的行为规律和神经机制研究	浙江工业大学
	面向取送货路径需求的多仓门越库系统车辆调度研究	广东工业大学
	动态环境下客户需求驱动的多 CODP 可演化生产建模、优化与应用	暨南大学
	融资约束影响专业大户创业的理论与实证研究	南京财经大学
	面向知识服务的知识库结构研究	南京大学
	支持非常规突发事件应急决策的模型形式化表示和模型组合方法研究	南京航空航天大学
	基于 CSSCI 的句法级汉英平行语料库构建及知识挖掘研究	南京农业大学
	新媒介影响下的突发事件不实信息传播机理与动态控制策略研究	上海理工大学
	基于语义的医学领域前沿知识发现及演化机制研究	中国医学科学院
	自行车微观行为机理及其非线性动力学建模研究	北京交通大学
	多粒度语言词计算理论及其在“情景-应对”型应急决策中的应用	南京信息工程大学
	面向价值共创的产品服务系统个性化配置研究	上海电机学院
	分布异构环境下基于多 Agent 的复杂供应链网络计算实验方法研究	浙江财经大学
	大型武器装备全寿命风险动态演化的多分辨率建模方法	中国人民解放军海军工程大学
	产学研联盟中企业联盟管理能力的形成及对创新的影响机理研究	武汉大学

AMiner

	专业虚拟社区与企业电子知识库间的知识转移研究：大数据与用户行为相结合的视角	西安理工大学
	心理健康众包系统的分析、建模与设计	香港城市大学深圳研究院
	专利文献的要素组合拓朴结构及向量空间语义表示与相似度计算研究	江苏大学
	基于 FrameNet 的中文评价词汇本体构建与观点挖掘研究	山西大学
	面向科技监测的实体识别与关系抽取研究	中国科学技术信息研究所
	化工园区事故情景演化重构及应急决策研究	中国石油大学（华东）
	面向多主体的应急预案体系语义模型研究	大连理工大学
	地学境本协同语义驱动的矿山企业生产知识服务模式构建研究	东北大学
	面向智能互联网络的实体信任关系理论研究	合肥工业大学
	面向场外核应急响应的集成化决策空间研究	南华大学
	考虑产品间协同效应的多产品组合采购问题研究	中国石油大学（北京）
	集成情境信息的协同物流柔性运作过程建模与分析	北京交通大学
	非常规突发事件中群体压力下非理性“羊群现象”从众行为的临机决策理论与方法	上海大学
	创新用户细分建模与需求表征应用研究	深圳大学
	投资者有限关注与证券市场监管：基于大数据和计算实验的方法	南京农业大学
	精神医学领域科研合作及研究前沿趋势预测	山西医科大学
	网络本体质量及适应性的评估研究	武汉大学
	基于本体的医学科研统计设计与统计分析方法选择研究	中国人民解放军第四军医大学
	基于语义网络的传统针灸概念体系表示及应用研究	中国中医科学院中医药信息研究所
	基于本体的中医诊疗信息模型构建研究	中国中医科学院中医临床基础医学研究所
	基于复杂网络理论的肺系病中医古籍本体构建与知识发现研究	广州中医药大学
	基于本体的我国心血管类药物不良反应通路预测方法研究	吉林大学
	基于动态联邦查询优化算法的医药知识库构建方法研究	中国科学院深圳先进技术研究院
	基于语义网络的肾系疾病古籍本体知识库构建及方证规律挖掘	广州中医药大学

AMiner

	基于本体与多智能体的预制构件供应链信息交互机制及决策仿真研究	上海大学
	基于知识本体的 EPC 项目知识治理模式研究	大连理工大学
	大规模知识图谱上相似节点查询技术研究	北京大学
	基于知识图谱的个性化 API 推荐及组合技术研究	中国科学院信息工程研究所
	基于模糊本体和情境感知技术的用户行为偏好分析及个性化推荐研究	齐鲁工业大学
	基于子图近似匹配的海量知识图谱分布式查询技术研究	东南大学
	多维情景本体构建与协同研究	北京交通大学
	基于非结构化文本的知识图谱扩充方法研究	中国科学院自动化研究所
	基于知识图谱的用户长尾需求建模研究	中山大学
	基于语义网的微生物多源异构数据整合关键技术研究	中国科学院微生物研究所
	大数据环境下基于生物医学本体的文献无监督分类方法研究	上海海事大学
	基于多源数据融合的基因本体扩展方法研究	西北工业大学
	基于迁移学习的地理领域知识图谱构建技术	武汉理工大学
	基于多标签语义本体的图像深度哈希学习新方法研究	西北工业大学
	基于层次化情感本体和深度学习的社交媒体视觉内容情感分析	郑州轻工业学院
	多源知识图谱下大规模语义数据的融合及检索方法	郑州轻工业学院
	基于生成对抗网络的释义文本生成研究	黑龙江工程学院
	异构分布式并行环境下超大规模 RDF 知识图谱数据管理技术研究	湖南大学
	基于知识库的学习路径优化方法研究	中国科学院软件研究所
	基于教育知识图谱的个性化学习路径自动生成研究	北京师范大学
	大规模知识图谱驱动的复杂产品方案创新设计方法及其应用	浙江大学
	地理知识图谱实体关联关系建模与补全方法研究	中国人民解放军信息工程大学
	面向舆情知识图谱的演化图数据事件查询与分析技术	辽宁大学
	基于教育知识图谱的在线协作学习交互分析关键技术研究	北京师范大学
	基于多源异构数据的知识图谱补全及验证关键技术研究	东华大学
	基于知识图谱技术和危险本体模型的危险辨识关键技术研究	华东理工大学
	基于知识图谱的问答系统的鲁棒性研究	上海财经大学

AMiner

	基于混合域嵌入式向量模型的知识图谱自动演进问题研究	西南财经大学
	基于表示学习的知识图谱近似查询方法研究	东南大学
	基于结构化数据表示学习的文本生成技术研究	哈尔滨工业大学
	面向医学知识图谱演进的图神经网络知识融合及补全方法研究	北京邮电大学
	基于知识图谱的业务流程模型多层次实体匹配研究	淮阴师范学院
	基于动态数据的灰色关联分析法构建中医体质动态变化规律知识图谱	成都中医药大学
	基于多数据源医学知识图谱的降压药物决策模型研究	中南大学
	大数据背景下基于知识图谱的电子产品回收政策环境效应与经济性分析方法	浙江大学
	基于深度学习与知识图谱相结合的水下三维场景感知研究	中国海洋大学
	矿床领域文本数据挖掘与知识图谱构建	中国地质大学（武汉）
	多路聚合的跨媒体模式分析与知识推理研究	郑州大学
	基于注意力机制的社交网络舆情引导文本生成方法与策略	西华大学
	非结构化数据中威胁本体构建、实体识别与关系抽取方法研究	四川大学
	面向知识图谱自然语言问答的结构化查询生成关键技术研究	复旦大学
	强化学习框架下的知识图谱构建技术研究	中国人民解放军国防科技大学
	本体驱动的地址数据空间语义建模与地址匹配方法	深圳大学
	基于科研创新知识图谱的专利推荐方法研究	中国人民大学
	基于知识图谱和社交网络的在线健康社区医患服务匹配研究	浙江工业大学
面上项目	量子逻辑和模糊逻辑的相关问题研究	湖南大学
	谓词逻辑与模型检验中的计量化理论	陕西师范大学
	基于对合否定的 SBL 公理化扩张系统的程度化推理及逻辑控制研究	延安大学
	基于形式化语义方法的模糊泛函分析理论及自然语言定性处理的逻辑基础	南京师范大学
	汉语句子理解中语义和句法整合的认知神经机制	华南师范大学
	基于自适应学习的农业领域本体建模理论与方法研究	安徽农业大学
	基于循证医学本体论的临床元数据语言研究	中国人民解放军第二军医大学
	基于领域本体方法构建中医概念信息模型的研究	中国中医科学院中医药信息研究所

AMiner

	梭菌糖代谢途径知识库的构建	中国科学院上海生命科学研究院
	微生物药物生物合成转化知识库的构建	武汉大学
	新生儿脓毒症诊断标志物的知识库构建及新型组合标志物的发现	苏州大学
	基于本体论的地理信息分类和多源空间数据融合研究	武汉大学
	人类视觉认知与多尺度遥感图像智能化处理方法研究	华南师范大学
	基于语义网的多源地学空间数据融合与挖掘研究	上海财经大学
	时空变化本体论分析研究	国家测绘地理信息局卫星测绘应用中心
	地理信息检索中语境的获取、推理及应用	中国科学院地理科学与资源研究所
	旅游地理本体与众包旅游信息的时空语义挖掘研究	河北师范大学
	基于空间认知常识的定性地理信息检索研究	武汉大学
	基于本体论的海洋流动场时空数据建模与可视化	山东科技大学
	基于语义网格的跨学科制造资源深度挖掘与服务发现研究	浙江财经大学
	基于语义网关的制造企业异构信息系统集成建模方法和实现技术	重庆大学
	东北地区既有住宅病理诊断及有效修复的辅助知识库研究	大连理工大学
	基于领域本体的煤矿安全数据融合方法及应用	安徽理工大学
	基于本体的 GPS 信息语义互操作方法研究	华中科技大学
	优化算法和知识库驱动的方案阶段建筑节能设计与技术工具原型研究	东南大学
	基于 fMRI 的个性化图像情感标注及其本体库研究	太原理工大学
	大规模本体的分块映射及相关评价方法研究	中南大学
	语义网本体不精确性模型研究	大连海事大学
	超协调缺省逻辑及其应用	北京大学
	面向认知网络的自律计算模型及评价方法研究	哈尔滨工程大学
	基于本体的深层网络数据集成方法研究	哈尔滨工程大学
	基于本体的 Deep Web 搜索技术	吉林大学
	基于认知语境的文本情感计算及其应用	大连理工大学
	大规模动态本体映射模型研究	清华大学
	事件本体模型与应用技术	上海大学
	认知模型驱动的海量中医医案知识获取技术研究	北京科技大学
	本体学习的粒计算模型与方法研究	南昌大学
	引入命题逻辑支持组合优化问题的求解——以图顶点染色问题为研究介质	华中科技大学

AMiner

	基于认知的动态波束形成技术研究	中国人民解放军理工大学
	消费者在线评论情感近似推理机制及其应用研究	大连理工大学
	情景演算中的关键推理技术及其应用研究	中山大学
	面向多知识主体不确定性推理的本体可信模型研究	同济大学
	海上立体搜救全局最优化模型与知识库支持方法研究	大连海事大学
	数据库支持的语义 Web 模糊本体管理关键技术的研究	东北大学
	时变目标动态证据推理及应用研究	西北工业大学
	基于观测角度的汉语名词性隐喻逻辑释义和评价方法研究	厦门大学
	面向 Deep Web 的大规模知识库自动构建方法研究	苏州大学
	动态时空推理研究	吉林大学
	面向本体映射的在线语义连接子研究	扬州大学
	粒计算拓展模型及多知识库集成信息融合研究	中南大学
	基于本体的临床路径建模与电子化方法研究	浙江大学
	基于流形学习和时序语义网挖掘的人体运动序列分析研究	大连理工大学
	基于语流音变认知机理建模的语音识别研究	北京语言大学
	扩展的模糊逻辑与基于蕴涵算子的 Rough 逻辑	上海海事大学
	基于本体的专利自动标引研究	北京信息科技大学
	基于约束松弛的概率图模型近似推理研究及在计算摄像学中的应用	清华大学
	基于概率推理求解命题逻辑可满足性问题的局部搜索技术研究	华中科技大学
	面向领域本体的多源异构数据聚合和语义标注关键技术研究	洛阳师范学院
	不确定性概念内涵与外延的双向认知计算理论模型与方法	重庆邮电大学
	基于形式概念分析的描述逻辑本体构建理论与方法	华南师范大学
	构建面向 Web 的、以实体为中心的知识库的关键技术研究	清华大学
	本体导向的大规模语义信息声明式抽取方法研究	武汉大学
	面向关系数据库知识发现的概率逻辑贝叶斯网络研究	吉林大学
	中文动态语义网构建技术研究	北京大学
	面向人的高级认知功能系统化研究的多维脑数据概念模型	北京工业大学
	描述逻辑中的本体融合方法研究	东南大学
	事件本体形式化方法中的几个重要问题	上海大学
	大规模 OWL 本体的分布式调试方法研究	东北大学
	语义 Web 中典型不确定知识的本体表示和融合的理论与方法研究	首都师范大学

AMiner

	具有推理能力的时空对象情境感知技术	南京航空航天大学
	基于动作概念的本体知识库及在文本处理上的应用	上海交通大学
	资源感知的程序逻辑理论及资源安全性推理	深圳大学
	汉语篇章框架语义关系网自动抽取及其语义推理	山西大学
	区分矛盾否定、对立否定和中介否定的模糊集 FScom 与模糊逻辑 FLcom 理论及其在典型的知识处理领域中应用的研究	江南大学
	异构环境下基于社交数据的大规模本体学习模型研究	清华大学
	缺失和不一致数据环境下描述逻辑查询回答的高效方法研究	广东外语外贸大学
	实际环境中基于联想记忆网络的增量式推理系统研究	南京大学
	基于短文本的知识库自动更新关键技术研究	北京理工大学
	本体匹配中的参数和策略调谐问题研究	东南大学
	面向大规模知识图谱的查询处理关键技术研究	复旦大学
	基于时空上下文数据的关联关系挖掘与推理技术研究	华中科技大学
	基于微博社区的知识图谱构建与分析	西华大学
	基于词向量表示的大规模知识图谱构建方法研究	中国人民大学
	多粒度信息的多值逻辑描述及其计量化知识推理	西安石油大学
	基于深度学习的面向海量科技文献的大规模高性能本体构建技术研究	东北大学
	不确定知识图谱中面向结构查询的众包清洗研究	华东师范大学
	大规模知识图谱的分布式表示学习、知识获取与推理应用	清华大学
	多智能体动作推理及高级控制的理论与技术研究	中山大学
	面向大规模知识图谱的弱监督中文实体关系抽取研究	华中师范大学
	稳定模型语义下不一致本体的存在规则语言修复理论研究与实现	中山大学
	不完全与不协调知识表示及推理	北京大学
	基于大规模知识库的问答系统关键技术研究	北京大学
	融合知识图谱的文本个性化推荐机制研究	北京理工大学
	语义 Web 模糊时空本体知识管理关键技术研究	东北大学
	基于图的并行 OWL 本体推理方法研究	东南大学
	基于情感知识库构建、领域适应和终身学习的社交媒体情感分析方法	南京理工大学
	复杂组群行为中动作元建模及深层交互关系模型推理关键问题研究	青岛科技大学
	基于知识图谱的在线个性化商品推荐技术研究	浙江大学
	多模异构生物特征深度融合、关联推理及其一致性语义挖掘关键技术研究	华侨大学
	时空知识图谱的表示模型与计算方法研究	浙江大学

AMiner

	基于贝叶斯本体的 Web 知识发现研究	北京航空航天大学
	基于本体论的企业默会知识转化与共享管理研究及应用	山东工商学院
	中文语境下基于模糊本体的用户在线评论的情感分析	同济大学
	诊疗本体自动构建方法与过程驱动的本体进化机制研究	上海交通大学
	基于生命周期成本理论的知识库构建及应用研究	中南大学
	基于领域本体的产品质量安全 Web 信息分析	北京航空航天大学
	支持智能化 BIM 技术的建筑业知识本体方法研究	天津大学
	基于模式识别的动态过程质量监控与诊断	郑州大学
	基于社会治理视角的公立医院社会评价路径及策略研究	杭州师范大学
	语义网络环境下数字图书馆资源多维度聚合与可视化研究	吉林大学
	面向知识服务的知识组织模式与应用研究	南京大学
	社会化媒体集成检索与语义分析方法研究	武汉大学
	数字图书馆社区的知识聚合与服务研究	武汉大学
	基于关联数据的信息聚合模型与实现研究	浙江大学
	基于海量数字资源的科研关系网络构建研究	中国科学技术信息研究所
	我国卫生决策支持系统理论与实证研究	中国医学科学院
	大型建设工程项目团队知识学习的机理及制度研究	长沙理工大学
	集装箱码头同贝同步装卸的智能调度优化与干扰管理方法研究	大连海事大学
	支持社会化创新的在线社区知识活动动力学研究	大连理工大学
	订单装配型供应链环境下的个性化产品配置及优化建模研究	东华大学
	二型模糊系统理论及其在知识个性化推荐中的应用	东南大学
	多项目运作环境下基于项目族的关键资源配置研究	湖南大学
	基于超网络的企业微博知识挖掘及整合方法研究	华南理工大学
	基于兴趣图谱的社会网络演化机制研究	上海理工大学
	在线评论对商家销售业绩的影响：情感分析视角	同济大学
	精细化科研管理信息系统研究	香港城市大学深圳研究院
	遥操作中空间认知过程及绩效的研究	中国航天员科研训练中心
	基于多方论据博弈的决策知识萃取与联合学习方法研究	中国人民解放军国防科学技术大学

AMiner

	制造企业逆向迁移、网络重构与高端知识获取：基于三阶段网络演进的视角	浙江工商大学
	企业间网络对突破性创新影响机制：基于知识基础配置与知识整合视角	华东理工大学
	成员异质性对创新网络中成员互动及创新绩效影响研究	哈尔滨理工大学
	基于演化本体的网络舆情自适应话题跟踪方法研究	南京航空航天大学
	新研究领域科学文献传播网络生长及对传播效果影响研究	南京理工大学
	我国产学研协同创新的理论与实证研究	上海交通大学
	基于复杂网络的中文文本语义相似度研究	西安电子科技大学
	大数据环境下基于领域知识获取与对齐的观点检索研究	中南财经政法大学
	基于神经网络和强化学习的车辆装配系统中的多载量小车实时调度方法	同济大学
	钢铁工业逆向供应链服务模块化及自适应匹配机制研究	武汉科技大学
	高技术服务创新网络平台企业主导能力的演化机理研究	湖南大学
	网络环境下服务系统的自主管理研究	华南农业大学
	区域基础教育均衡发展机理与发展模式研究	河北师范大学
	大数据环境下多媒体网络舆情信息的语义识别与危机响应研究	吉林大学
	技术范式转换预警的理论与方法	南京工业大学
	面向词汇功能的学术文本语义识别与知识图谱构建	武汉大学
	大数据环境下基于特征本体学习的无监督文本分类方法研究	湖北工业大学
	基于社会商务知识图谱的信息流广告投放关键技术研究	西南财经大学
	嵌入式知识服务驱动下的领域多维知识库构建	吉林大学
	基于个性化知识地图的交互式信息检索系统研究—从用户认知的角度	武汉大学
	基于文本逻辑主题结构的数字出版内容重组研究	武汉大学
	基于自动问答系统的中医诊疗认知模式研究	北京中医药大学
	面向临床指南计算机应用的医学信息表达本体及元模型构建	中国人民解放军第四军医大学
	基于多尺度非完整数据的中医个体知识图谱构建及其诊疗规律研究	中国中医科学院中医临床基础医学研究所
	基于多源异构数据的知识图谱构建、推理与问答研究	北京航空航天大学

AMiner

	基于本体学习与本体映射的组织异构数据融合方法研究	福州大学
	基于本体的糖尿病肾病防诊治精准决策支持研究	浙江大学
	概率推理的眶额叶神经环路及机制	中国科学院上海生命科学研究院
	基于知识图谱架构的舰船编队环境态势描述技术研究	哈尔滨工程大学
	语义网的形式化查询的生成方法研究	南京大学
	面向互动语言场景的类量子语言模型关键理论和技术研究	天津大学
	基于隐含知识挖掘与时间敏感的知识图谱补全关键技术研究	北京大学
	基于知识图谱的突发事件预警预测技术	国家计算机网络与信息安全管理中心
	基于众包技术的知识图谱精化研究	华东师范大学
	深度形状语义网络表示机理及其结构保留机制与判别性研究	重庆大学
	基于概率推理的图像特征匹配统一理论模型及其应用研究	武汉大学
	基于深度学习的富媒体本体学习模型研究	清华大学
	面向文本生成的生成式对抗网络算法与理论研究	中国科学院计算技术研究所
	语义 Web 知识库补全关键技术研究	北京师范大学
	基于逻辑规则和表示学习的知识图谱关系推理方法与应用研究	电子科技大学
	基于本体的地理空间模型自动数据匹配方法研究	中国科学院地理科学与资源研究所
	基于知识图谱的社交媒体大数据主题发现与趋势预测关键问题研究	同济大学
	基于本体的复杂产品功能语义表示与智能分解研究	丽水学院
	知识图谱的多层网络耦合分析理论及其应用研究	上海财经大学
	XBRL 领域本体的自动构建研究	暨南大学
	基于知识库和深度学习的生物医学实体关系抽取研究	大连理工大学
	基于知识图谱的农业大数据碎片化知识发现方法研究	安徽农业大学
	知识图谱辅助的垂直领域自动人机对话系统框架研究	北京大学
	面向信息交互的中国临床标准药物知识库构建与评测研究	南通大学
	利用知识图谱进行生命组学数据知识发现的关键技术研究	中国人民解放军军事医学科学院
	基于知识图谱的轨道交通工程设计自动合规性检查关键技术	西安理工大学

AMiner

	融合深度学习的水电机组故障知识图谱构建与不确定推理诊断	华中科技大学
	面向情感交互的人机对话文本生成技术研究	东北大学
	异质网络环境下动态知识图谱构建技术研究	吉林大学
	基于词汇语义网络的中文深层语义分析	北京语言大学
	恶意代码知识表示与知识库构建方法研究	哈尔滨工业大学
	高精度稳定的黑盒变分推理关键问题研究	吉林大学
	基于噪音数据的中文知识图谱增量式构建研究	苏州大学
	基于问题预测和文本生成的阅读理解关键技术研究	北京大学
	基于溯因推理的知识图谱补全结果解释机制研究	广东外语外贸大学
	基于深度学习的数据-文本生成技术研究	北京大学
	公开信息、知识图谱与交易优化——基于大数据的视角	武汉大学
	群智制造环境下大规模知识图谱构建与可信服务组合优化研究	浙江财经大学
	稀疏样本条件下的领域知识图谱获取与利用研究	中国人民解放军国防科学技术大学
	基于大规模开放科学知识图谱的学科新兴趋势探测研究	武汉大学
	基于多模态机器学习的智能医疗决策知识推理研究	南开大学
	基于知识图谱的按专业招生高考志愿智能推荐方法研究	湖北大学
	多目标协同下语义化扩展的知识图谱补全机理及应用模式	大连海事大学
	面向知识图谱的规则学习和推理研究	天津大学
	基于知识图谱的高质量知识表示关键技术研究	北京理工大学
	基于跨语言信息的大规模稀缺资源知识库自动构建技术研究	哈尔滨工业大学
	面向机器人个性化服务的物品引擎式环境知识图谱构建机制	山东大学
	基于表示学习和生成对抗网络的藏文知识库问答关键技术研究	中央民族大学
	基于大规模知识库集成与多路径表示学习的开放域知识库问答研究	华中师范大学
	基于知识图谱的云制造服务跨界融合关键技术研究	天津大学
	尿液生物标志物知识图谱的建立及其在新生物标志物预测中的应用	北京蛋白质组研究中心
	情绪感知与知识增强的文本生成技术研究	中国科学院信息工程研究所
	“文本-地图”结合的地理知识图谱构建方法	南京师范大学

AMiner

	基于深度学习的学术全文本知识图谱构建及检索研究	南京理工大学
	云制造模式下基于混合模型与深度强化学习的知识图谱构建及其数据治理体系研究	浙江工业大学
	双曲空间中知识图谱联合文本的动态表征学习	中国地质大学（武汉）
	面向多模态智能问答的表达学习及知识推理关键技术研究	中山大学
	陕西传统村落景观的地方性知识图谱研究	西安建筑科技大学
	知识图谱与深度学习相融合的无人机 LiDAR 点云典型地物解译方法	集美大学
	基于词汇语义网络的语义表征与语义整合的认知神经机制研究	华东师范大学
	大规模知识图谱分布式存储与查询关键技术研究	天津大学
	综合减灾知识图谱构建方法研究	武汉大学
	基于建筑知识图谱的近代侨乡建筑形制与源流研究 -以跨境交流视野下岭南地区侨乡为例	中山大学
联合基金项目	用系统生物学方法对 miRNAs 介导的癌细胞死亡网络的研究	河南工业大学
	小麦芒长抑制基因 B1 的精细定位及其调控芒发育相关基因的分离	河南科技大学
	维吾尔族早发糖尿病 MODY 家系的血清和尿液蛋白质生物标志研究	新疆医科大学
	整体结构件光学测量工艺理论数值分析方法研究	郑州航空工业管理学院
	分布式模块化产品架构分类建模与技术状态控制研究	郑州轻工业学院
	高压电场下聚合物液体射流微观结构演变规律的同步辐射 SAXS 原位研究	中国科学院上海应用物理研究所
	面向军事情报的多媒体大数据分析与展示	中国科学院自动化研究所
	面向大数据的粒计算理论与方法	山西大学
	多尺度层次下的循环流化床中甲醇定向转化制烯烃过程的化学工程研究	上海交通大学
	面向智慧城市的大规模数据计算理论和关键技术	桂林电子科技大学
	面向甲骨学知识图谱的实体发现及语义关系挖掘研究	安阳师范学院
	复杂机电系统贝叶斯网络性能模型结构学习优化	河南理工大学
	高性能集成橡胶 SIBR 的合成与应用	大连理工大学
	造纸生产过程的分布式协同控制与智能优化	上海交通大学
	面向化工大数据的双极分析理论与方法研究	南京信息工程大学
	基于文本大数据的无载体信息隐藏研究	南京信息工程大学

AMiner

	维吾尔语汉语语音翻译系统关键技术研究	新疆大学
	基于溶剂-应力开裂的空天使役 PC/膜层透明件环境可靠性诊断及机理研究	郑州大学
	N-双相掺杂 Li4Ti5O12@C 介孔纳米片状材料可控制备及储钠机理	常州大学
	复杂多障碍环境中的物流服务机器人设计、控制、调度关键技术	中国科学技术大学
	基于跨媒体网络大数据的民航突发事件应急决策语义服务关键技术研究	中国民航大学
	缺乏载体先验知识的隐写分析模型与方法研究	中国科学院信息工程研究所
	大数据和社会场景下的互动语言意图深度理解基础理论和关键技术研究	天津大学
	基于移动大数据的特异群组挖掘与行为预测	复旦大学
	社会媒体大数据感知与分析关键技术研究	北京航空航天大学
	面向大数据的国际主题事件推演与风险预警研究	北京航空航天大学
	政府治理大数据行为知识图谱关键技术研究	中国人民大学
	面向开放域数据的情报知识图谱构建技术研究	北京邮电大学
	面向特定领域的知识图谱构建与应用关键技术研究	中国信息安全测评中心
国家基础科学人才培养基金	中国地质大学（武汉）地质学基地科研训练及科研能力提高项目	中国地质大学（武汉）
	广州中医药大学医学理科基地人才培养支撑条件建设项目	广州中医药大学
国际(地区)合作与交流项目	面向汉英双语的认知情感网和多模态话题分析技术研究	清华大学
	多语言多模态知识图谱关键技术研究及其应用	清华大学
	大数据环境下的知识组织与服务创新研究	武汉大学
地区科学基金项目	面向装配规划的领域知识表示及推理研究	桂林电子科技大学
	面向数据密集型计算的概率图模型构建与推理	云南大学
	基于社会性标注及本体的深层网语义搜索方法研究	桂林电子科技大学
	基于模糊 Petri 网的复杂知识系统推理、学习及应用研究	新疆大学
	智能机器人自主知识与自主行动推理研究	广西师范大学
	基于描述逻辑及符号算法的事例相似性研究	桂林电子科技大学
	本体学习的认知模型及其算法研究	云南大学
	面向语义 Web 的行动表示和推理研究	桂林电子科技大学

AMiner

	基于领域本体与技术规范的多代理分布式智能教学系统互操作模型研究	广西师范大学
	蒙古语词汇语义网研究	内蒙古师范大学
	基于概率本体的 CPS 入侵检测方法研究	桂林电子科技大学
	基于本体及推理机的构件化软件演化信息获取及度量技术研究	江西师范大学
	面向事件分析的信息意图检测、建模与群体意图推理技术研究	石河子大学
	本体的解释诊断理论研究	黔南民族师范学院
	安全领域本体构建的动态描述逻辑研究及多步攻击挖掘	昆明理工大学
	藏医诊疗知识库系统关键技术研究	青海大学
	形式化描述高水平智能机器人的逻辑基础研究	广西师范大学
	基于规则和统计的蒙古语多层级文本查错及纠错推理模型研究	内蒙古大学
	基于领域本体的蒙古文数字资源整合机制研究	内蒙古大学
	基于移动社会网络的旅游信息服务可信性分析及其应用研究	海南大学
	全面开放格局下中国银行业效率机制研究：基于金融稳定性和 FDI 的视角	海南大学
	泛在环境下基于情境历史和兴趣社区的个性化信息推荐模型与实现	江西财经大学
	农业数字防灾减灾资源规划机理分析与系统实现：基于 EA 和 Ontology 的研究	南昌大学
	单病种质量管理体系的数字化模型研究	云南省第一人民医院
	流程监控与评估中多元数据整合研究	内蒙古大学
	基于交通流量概率推理的不规则交叉口交通信号配时参数优化研究	延边大学
	可操作行为规则挖掘的建模与算法研究	大理大学
	基于本体的公差指标自动生成方法研究	桂林电子科技大学
	图模式匹配的软约束模型及其符号推理研究	桂林电子科技大学
	构建传统蒙古文知识图谱关键技术研究	内蒙古大学
	面向语言处理的藏语概念框架语义知识库构建研究	西北民族大学
	东盟多语种事件知识图谱构建与推理方法研究	桂林电子科技大学
	基于知识图谱的农产品价值链信息融合研究	贵州财经大学
	基于多维度连续情感空间的文本生成技术研究	云南大学

AMiner

	基于视觉-语义多模态知识图谱的情感嵌入式知性对话生成算法研究	江西师范大学
	集成注意机制与知识推理的图像描述生成	广西师范大学
	基于知识图谱的云南植物领域垂直搜索系统研究与应用	云南农业大学
	面向旅游领域的蒙古文知识图谱构建关键技术研究	内蒙古大学
	舆情本体概念间非分类关系抽取方法研究	新疆大学

AMiner

顾 问：唐 杰　杨红霞（阿里集团—新零售智能引擎事业群）

编 写：徐 菁

数 据：赵慧军

封面设计：徐 菁

校对排版：万文正　李可欣

资料协助：高云鹏　吴蓓　赵杨奥

更多 AI TR 系列报告，敬请查看官方网址 http://reports.aminer.cn/

如果您对科技写作有兴趣和经验，或有商务合作需求，

欢迎联系我们：reports@aminer.cn



考研资料：数学、英语、政治、管综、西综、法硕等（整合各大机构）
英语类：四六级万词班专四专八雅思等
财经类：初级会计、中级会计、注册会计、高级会计、税务师、会计实操、证券从业、基金从业、资产评估、初级审
公务员：国考、省考、事业单位、军队文职、三支一扶微信 2270291360
银行：银行招聘、笔试、面试
教师资格：小学、中学、教师招聘面试
建筑：一建、二建、消防、造价
法考：主观题、客观题
多平台网课：涵盖职场、办公技能、编程、文案写作、情感心理、穿搭技巧、理财投资健身减肥摄影技术等优质内容
精选资料：Excel 教程、PPT 模板、简历模板、PS 教程、PPT 教程、素描、烹饪、小语种、CAD 教程、PR 教程、UI 课程、自媒体、写作、计算机二级、钢琴、Python、书法、吉他、kindle 电子书、演讲.....持续更新中...
押题：提供考前冲刺押题（初级会计、中级会计、注册会计、一建、二建、教资、四六级、证券、基金、期货等等），麻麻再也不用担心我考不过了。
资料领取微信：1131084518
行业报告：20000 份+持续更新

英语四六级备考资料	计算机二级备考资料	150 所高校考研专业课资料
两小时搞定毛概马原思修近代史纲	高数(微积分)+线性代数+概率论	素描 0 基础入门教程
教师资格证全套备考资料	普通话考试资料礼包	书法教程微信 2270291360
大学生英语竞赛备考资料	大学生数学竞赛备考资料	1000 份各行业营销策划方案合集
挑战杯/创青春/互联网+竞赛资料	电子设计竞赛必备资料	街舞 0 基础入门教程
托福雅思备考资料	大学物理学科攻略合集	动漫自学教程
SCI 最全写作攻略	TEM4/TEM8 专四专八备考资料	教师资格证面试试讲万能模板
360 份精美简历模板	数学建模 0 基础从入门到精通	100 套快闪 PPT 模板
Vlog 制作最全攻略	超强 PR 模板	42 套卡通风 PPT 模板
PS 零基础教程微信 1131084518	PS 高级技能教程	63 套酷炫科技 PPT 模板
好用到极致的 PPT 素材	128 套中国风 PPT 模板	32 套 MBE 风格 PPT 模板
327 套水彩风 PPT 模板	295 套手绘风 PPT 模板	54 套毕业答辩专属 PPT
196 套日系和风 PPT 模板	82 套文艺清新 PPT 模板	57 套思维导图 PPT 模板
163 套学术答辩 PPT 模板	53 套北欧风 PPT 模板	34 套温暖治愈系 PPT 模板
118 套国潮风 PPT 模板	30 套仙系古风 PPT	126 套黑板风 PPT 模板
114 套星空风格 PPT 模板	192 套欧美商务风 PPT 模板	42 套绚丽晕染风 PPT
50 套精美 INS 风 PPT 模板	56 套水墨风 PPT 模板	137 套清爽夏日风 PPT 模板
98 套森系 PPT 模板	25 套简约通用 PPT 模板	记忆力训练教程
300 套教学说课 PPT 模板	123 套医学护理 PPT 模板	AE 动态模板微信 2270291360
毕业论文资料礼包	教师资格证重点笔记+易错题集	表情包制作教程
吉他自学教程（送 6000 谱）	钢琴自学教程（送 1000 谱）	区块链从入门到精通资料
2000 部 TED 演讲视频合集	Excel 从入门到精通自学教程	单片机教程
230 套可视化 Excel 模板	1000 款 PR 预设+音效	1000 份实习报告模板
手绘自学教程微信 1131084518	单反从入门到精通教程	人力资源管理师备考资料
英语口语自学攻略	粤语 0 基础从入门到精通教程	证券从业资格证备考资料
日语自学教程	韩语自学教程	PHP 从入门到精通教程
法语学习资料	西班牙语学习资料	炒股+投资理财从入门到精通教程
全国翻译专业资格考试备考资料	BEC 初级+中级+高级全套备考资料	大数据学习资料
SPSS 自学必备教程	Origin 自学必备教程	会计实操资料
LaTeX 全套教程+模板	EndNote 教程+模板	小提琴 0 基础入门自学教程
GRE 超全备考资料	200 份医学习题合集	司考备考资料

思维导图教程＋素材	20000 份行业研究报告	一级造价工程师备考资料

上万 GB 教学资料（均全套,非杂乱）免费领取微信 1131084518

《闪电式百万富翁》实战版+升级版

易经+道德经+易学名师全集+风水学+算命学+起名+++等等（全套 1000 多 GB）

心理学+NLP 教练技术+精神分析+亲子家庭教育+催眠+++等等（更新超 2000GB）

大学-已更新至 9333 个课程+高中+初中+小学-全套资料（超过 2 万 GB）

陈安之	曾仕强	马云	杜云生	翟鸿燊	刘一秒	俞凌雄
王健林	余世维	雷军	周文强	安东尼罗宾	董明珠	李嘉诚
徐鹤宁	冯晓强	李践	刘克亚	罗伯特清崎	戴志强	李伟贤
苏引华	史玉柱	李强	俞敏洪	杰亚伯拉罕	周鸿祎	唐骏
梁凯恩	陈永亮	傅佩荣	贾长松	易发久	李彦宏	湖畔大学
李开复	聂枭	悟空道场	魏星	姬剑晶	其他名师全集	其他资料下载
王兴	王智华	智多星	陈文强	周导		微信 2270291360
泡妞	撩汉	泡仔	房中术	性福课	驻容术	泡妞撩汉性福合集

注：太多了，无法全部一一列出。。。

全套专题系列【微信 1131084518】

记忆力训练	形象礼仪	健康养生	企业管理	沟通技巧
演讲与口才	经理修炼	MBA 讲座	时间管理	战略经营
企业文化	销售心理	管理素质	国学讲座	执行力
团队管理	领导艺术	员工激励	潜能激发	谈判技巧
绩效管理	薪酬管理	43 份直销制度	电话销售	人力管理
客户服务	创业指南	市场营销	餐饮管理	保险讲座
品牌营销	酒店管理	汽车 4S 店	众筹资料	销售技巧

兴趣爱好：	钓鱼教程	魔术教学	炒股教学	美术教学	书法教学	
音乐乐器：	萨克斯教学	电子琴教学	小提琴	古筝教学	钢琴教学	
	吉他教学					
体育运动：	篮球教学	足球教学	羽毛球教学	乒乓球教学	太极拳教学	
	围棋教学	高尔夫球				
生活实用：	插花教学	茶艺-茶道	唱歌教学	单反相机摄影	毛线编织	
	小吃+美食					暗号：666
语言学习：	英语					免费领取资料微信
电脑 IT：	办公 office	PS 美工教学				1131084518

微信1131084518
撩汉liaohan.net
最好资源zuihaoziyuan.com



如果群里报告过期

请加微信联系我索取最新

1、每日微信群内分享7+最新重磅报告;

2、每日分享当日华尔街日报、金融时报

3、如果看到群里报告过期了，请扫码联系

4、行研报告均为公开版，权利归原作者所有，仅分发做内部学习

扫一扫二维码后台回复 加群

加入“研究报告”微信报告群。

